

NBER WORKING PAPER SERIES

PAYMENT SYSTEMS AND INTERCHANGE FEES

Richard Schmalensee

Working Paper 8256  
<http://www.nber.org/papers/w8256>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
April 2001

I am indebted to National Economic Research Associates and Visa U.S.A. for financial support. Howard Chang, David Evans, James Hunter, Jean Tirole, Kyle Bagwell, Robert Porter, anonymous referees, and, especially, Bernard Reddy provided valuable comments on earlier versions of this paper. I owe a special debt to the late William F. Baxter, who attracted me to this problem. I retain sole responsibility for imperfections and opinions. The views expressed herein are those of the author and not necessarily those of the National Bureau of Economic Research.

© 2001 by Richard Schmalensee. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Payment Systems and Interchange Fees  
Richard Schmalensee  
NBER Working Paper No. 8256  
April 2001  
JEL No. L

**ABSTRACT**

In a typical bank credit card transaction, the merchant's bank pays an *interchange fee*, collectively determined by all participating banks, to the cardholder's bank. This paper shows how the interchange fee balances charges between cardholders and merchants under imperfect competition. The privately optimal fee depends mainly on *differences* between cardholders' and merchants' banks, not their collective market power. In a non-extreme case, the profit-maximizing interchange fee also maximizes total output and producers' plus consumers' surplus. There is no economic basis for favoring proprietary payment systems, which do not need interchange fees to balance charges, over the cooperative bank card systems.

Richard Schmalensee  
Sloan School of Management  
Massachusetts Institute of Technology  
Cambridge, MA 02138  
and NBER

## I. INTRODUCTION

In a bank credit card transaction, the bank that has issued the card to the consumer is called the issuing bank or *issuer*, and the bank that processes the transaction for the merchant is called the acquiring bank or *acquirer*. When the issuer and acquirer are different, the acquirer pays the issuer an *interchange fee*, set collectively by the banks that belong to the system. Interchange fees differ among transactions of various sorts; in recent years, interchange fees in the Visa and MasterCard systems have averaged between one and two percent of transaction value. Changes in interchange fees generally affect *merchant discounts*, the fees paid by merchants to acquiring banks for processing credit card transactions. In the U.S., where acquiring is highly competitive, changes in interchange fees lead to roughly equal changes in merchant discounts.

In the U.S., collective determination of interchange fees by competing banks was found to be legal in the 1984 *Nabanco* decision.<sup>1</sup> This decision rested in part on the analysis presented by William Baxter [1983]. Baxter argued that because payment system volume is determined by the actions of both issuers and acquirers, and because interchange fees merely shift costs between these two sides of the system, collective determination of the interchange fee is not ordinary anti-competitive price-fixing. He showed that under perfect competition among issuers and among acquirers, the socially optimal interchange fee is generally non-zero.

---

<sup>1</sup> National Bancard Corp. v. Visa U.S.A., Inc., 596 F. Supp. 1231 (S.D. Fla. 1984), *aff'd*, 779 F.2d 592 (11<sup>th</sup> Cir. 1986), *cert. denied*, 479 U.S. 923 (1986). For alternative views of the economics of this case, see Carlton and Frankel [1995a, 1995b], Evans and Schmalensee [1995], Frankel [1998], and Balto [2000].

Collective determination of interchange fees has recently come under renewed attack, particularly in Australia and the European Union.<sup>2</sup> One important element of this attack is the charge that because interchange fees are set to maximize profits of payment system members, rather than social welfare, it is appropriate to treat collective determination of interchange fees as cartel behavior.<sup>3</sup> Some have argued that collective determination should simply be banned, though it is not obvious whether bilateral negotiations between issuers and acquirers would lead on average to lower or (as Small and Wright [2000] argue) higher fees. Others (e.g., Balto [2000]) have argued that interchange fees should be set to zero by fiat or determined by regulators on the basis of system-related costs incurred by issuers and acquirers.

This paper analyzes the economic role played by the interchange fee in a payment system composed of profit-seeking, imperfectly competitive firms.<sup>4</sup> Two facts that served to

---

<sup>2</sup> This practice has recently been criticized by the Reserve Bank of Australia and the Australian Competition and Consumer Commission [2000] and has been formally challenged by the Competition Directorate-General of the European Commission [2000]. (See also Hehir [2000].)

<sup>3</sup> See Balto [2000] and the references cited in the preceding footnote. In addition, Frankel [1998] and others have argued that by increasing the merchant discount, a positive interchange fee magnifies the distortion created because merchants are prevented (by credit card system rules and/or by transactions costs) from imposing surcharges on customers who use credit cards, even though they are more expensive to serve than customers who use cash or checks. (For a response to Frankel [1998], see Evans and Chang [2000].) Schwartz and Vincent [2000] have recently formalized this critique of merchant discounts, while in the model of Rochet and Tirole [2000], merchant surcharging can increase or decrease welfare. In a model in which credit cards serve to increase total transaction volume by enhancing liquidity, Wright [2000] finds that merchant surcharging tends to reduce welfare by reducing cardholding. These analyses each rest on different simplifying assumptions to permit tractable modeling of consumers and retailers (all neglect search and search costs, for instance, which are central to some models of retailing). All neglect the facts that cash and checks are regulated and subsidized and that their costs to merchants generally differ. In light of all the departures from first-best optimality in this context, the theory of the second-best suggests that regulating card system merchant discounts will raise welfare only by chance. There is even less reason to think that welfare would be increased by regulating only the discounts of the bank card systems (via attacks on interchange fees) and not those of the proprietary systems. I will, in any case, neglect these issues for simplicity in what follows.

<sup>4</sup> Unless competition is at least slightly imperfect, it is hard to model choice of the interchange fee at the system level. Evans and Schmalensee [1995, pp. 899-901] show that in a perfectly competitive world with no frictions, *any* interchange fee is consistent with a zero-profit market equilibrium. In such a world, if merchants cannot give a discount for cash or impose a surcharge for credit purchases, a single interchange fee, determined by costs and equal to zero only by chance, is consistent with market equilibrium.

motivate this work suggest that this role is quite unusual. First, ATM (Automatic Teller Machine) and debit card networks also generally set interchange fees collectively, but in some of these networks fees flow in the opposite direction – from issuers to acquirers.<sup>5</sup> A general analysis must thus be consistent in principle with interchange flows in either direction. Second, in the U.S., because American Express has always served as its own exclusive issuer and acquirer, it has nothing corresponding directly to the interchange fees of the Visa and MasterCard systems. Nonetheless, even though it has been smaller than both these systems in recent years, it has generally charged merchant discounts substantially *above* the Visa and MasterCard averages.<sup>6</sup>

The key assumption of the analysis here is that the value of a payment system to issuers is affected by the behavior of acquirers and vice versa. This network externality can only be addressed at the system level, and we show that the interchange fee provides a simple, though imperfect, tool for addressing it. The main economic role of the interchange fee is not to exploit the system's market power; it is rather to shift costs between issuers and acquirers and thus to shift charges between merchants and consumers to enhance the value of the payment system as a whole to its owners.<sup>7</sup> The sign and magnitude of the value-maximizing interchange

---

<sup>5</sup> For debit card networks, see Faulkner & Gray [1999, pp. 22-26] and Reserve Bank of Australia and Australian Competition and Consumer Commission [2000]. When a customer pays with a check, no party pays anything like an interchange fee. But this zero-fee regime was produced by the Federal Reserve, not unregulated market forces: see, e.g., Spahr [1926], Jessup [1967], Baxter [1983], and Frankel [1998].

<sup>6</sup> See Evans and Schmalensee [1999, chs. 6 and 8]. Discover/Novus operates as a proprietary system like American Express. It has generally charged lower merchant discounts than the bank card systems, though on average its discounts have exceeded the markups charged by bank card acquirers over the bank card systems' interchange fees.

<sup>7</sup> Particularly in the early years of the bank credit card systems, most banks functioned as both issuers and acquirers. The basic externality on which this analysis rests is still present, however, as long as the profits any particular bank earns from its issuing (acquiring) operations is affected by the actions of other banks' acquiring (issuing) operations.

fee depend on the system's objectives, on differences in costs and in demand elasticities of issuers and acquirers, differences in the intensity of competition on the two sides of the system, and, in general, on differences in spillover effects between them.

Under imperfect competition, no matter how vigorous, one would not expect the interchange fee (or any other price) to be chosen in a socially optimal fashion. It is thus remarkable that under non-extreme assumptions, the privately optimal interchange fee *is* also socially optimal: I show below that it maximizes both total system output and a conventional Marshallian measure of social welfare. More generally, in deciding whether collective determination of the interchange fee should be treated like ordinary cartel price-fixing, the key question is whether collective fee setting, like ordinary price-fixing, is generally used to increase profit by reducing output. The answer is clear: it is not. The privately optimal fee may be above or below the socially optimal fee, and the difference does not turn on the level of market power.

In a paper complementary to this one, Rochet and Tirole [2000] assume perfect competition among credit card acquirers and imperfect competition among issuers and retailers. They explicitly model the retail sector, allowing for strategic behavior, and simplify by assuming identical retailers. This simplification enables them to derive welfare measures from fundamental cost and preference assumptions. Rochet and Tirole focus on equilibria in which all retailers accept credit cards, while an important feature of the analysis here is that retailer acceptance varies among equilibria. Thus the Rochet-Tirole setup facilitates rigorous welfare analysis, while the assumptions made here facilitate exploration of the balancing role of the interchange fee. Consistent with the results obtained here, Rochet and Tirole find that the

profit-maximizing interchange fee never reduces the output of credit card services below the efficient level.

## II. BASIC ASSUMPTIONS

For simplicity, the exposition that follows concentrates on bank credit card systems, though the basic analysis applies more generally. Bank credit card systems are operated on a *cooperative* basis: they pass interchange fees through from acquirers to issuers, and they pay no dividends to the banks that own them.<sup>8</sup> In contrast, *proprietary* systems like American Express earn profits at the system level, whether they are *unitary* (and do all issuing and acquiring themselves) or *non-unitary* (and contract with others to do some issuing and/or acquiring). We explore some implications of these alternative structures at the end of Section IV.

Because the volume of transactions in any particular bank card system is determined by the interaction of consumers' decisions to use the card and merchants' decisions to accept it, actions of acquirers impose external effects on issuers and vice versa. Any particular card brand is more valuable to consumers the more merchants they expect to accept it, and accepting any particular card is more valuable to merchants the more consumers they expect to carry and use it. Finally, any given volume of transactions can in general be produced by an infinite number of combinations of household and merchant activity, and thus of acquirers' and issuers' efforts to stimulate demand.

---

<sup>8</sup> See, generally, Evans and Schmalensee [1993; 1999]. In contrast, some ATM and debit card networks are proprietary and earn significant profits for their owners (Kim [1998], Faulkner & Gray [1999]). Thus some of these networks impose charges on their members that exceed system-level costs, while the bank card systems do not. I explore some consequences of this alternative, proprietary regime at the end of Section IV. In all that follows I assume system-level costs to be zero for simplicity; they are in fact small relative to interchange fees.

Let  $Q^T$  be the value of transactions — each assumed for simplicity to have the same monetary value — on a bank card system. It is useful to begin with the simplest case of *bilateral monopoly*: a single issuer and a single acquirer. A convenient demand structure that illustrates the key system-level features discussed just above is the following:

$$(1) \quad Q^T = Q^m(P^a)Q^c(P^i).$$

The quantity  $Q^m$  reflects merchants' willingness to accept cards; it is a decreasing function of the per-transaction price,  $P^a$ , that is fixed by the acquirer and that corresponds to the merchant discount charged by all payment card systems. Similarly,  $Q^c$  reflects consumers' willingness to carry and use cards; it is a decreasing function of  $P^i$ , the effective per-transaction price charged by the issuer to consumers.  $P^i$  could take the form of average interest payments on outstanding account balances.<sup>9</sup> I refer to  $Q^m$  and  $Q^c$  as *partial demands* in what follows.<sup>10</sup>

One useful way to think of this demand structure is as follows. Consumers' desired level of transactions volume is given by (1) as a function of  $P^i$ , with  $Q^m$  treated by consumers as exogenous. Here  $Q^m$  embodies *network effects*: the lower is  $P^a$ , all else equal, the greater is merchants' aggregate willingness to accept cards, thus the higher is  $Q^m$ ; and the higher is  $Q^m$ , the greater the volume of card transactions desired by consumers. Similarly, merchants' demand for transactions is given by (1) as a function of  $P^a$ , with  $Q^c$  treated by merchants as

---

<sup>9</sup> In fact, total costs to both consumers and merchants may have fixed components – including annual fees, terminal installation costs, and transactions costs of dealing with an issuer or acquirer. With non-trivial fixed costs, expected per-transaction cost depends on frequency of use, which for consumers depends on expected merchant acceptance and for merchants depends on expected consumer use. I simplify by following the relevant literature and assuming that these sorts of fixed costs can be neglected in equilibrium because acceptance and use expectations are fulfilled.

<sup>10</sup> This specification of network effects is, of course, somewhat restrictive. Appendix B analyzes a model in which partial demand functions are linear, and each partial demand function exhibits network effects directly by being an increasing function of the expected partial demand on the other side of the system. This generalization complicates the analysis of the linear case considerably but does not change any fundamental conclusions.



exogenous. On this side of the system,  $Q^c$  embodies network effects. In equilibrium the transactions volumes desired by both consumers and merchants equal the actual volume.

On this interpretation, one can use equation (1) to derive Marshallian partial equilibrium (consumers' plus producers' surplus) welfare functions for consumers and merchants. Treating  $Q^m$  as a constant, for instance, solve (1) for  $P^i$  and integrate under the resulting demand curve to obtain

$$(2a) \quad U^c(Q^T; Q^m) = \int_0^{Q^T} Q^{c-1}(x/Q^m) dx,$$

where  $Q^{c-1}$  is the inverse of the  $Q^c$  partial demand function. Similarly, on the merchant/acquirer side of the system,

$$(2a) \quad U^m(Q^T; Q^c) = \int_0^{Q^T} Q^{m-1}(x/Q^c) dx,$$

where  $Q^{m-1}$  is the inverse of the  $Q^m$  partial demand function. Then if  $C^a$  and  $C^i$  are the acquirer's and issuer's constant per-transaction costs, respectively, the corresponding Marshallian welfare measure is given by

$$(3) \quad W = U^c(Q^T; Q^m) + U^m(Q^T; Q^c) - (C^a + C^i)Q^T.$$

This measure, of course, does not reflect the facts that merchants are not final consumers and that competition among merchants is likely to be imperfect. Nonetheless, given the popularity of Marshallian welfare analysis in a variety of policy settings — including, in particular, the analysis of regulated prices — this measure provides a potentially interesting benchmark here.

Continuing with the bilateral monopoly case for simplicity, if  $T$  is the per-transaction interchange fee, the acquirer's profit is

$$(4a) \quad \Pi^a = [Q^m(P^a)Q^c](P^a - C^a - T),$$

where we adopt the convention that the interchange fee is positive when it is paid (as in actual bank credit card systems) from acquirers to issuers.<sup>11</sup> Similarly, because the interchange fee is simply transferred by the system from the acquirer to the issuer in cooperative payment systems, the issuer's profit is given by<sup>12</sup>

$$(4b) \quad \Pi^i = [Q^c(P^i)Q^m](P^i - C^i + T).$$

Equations (4) show that if it were possible to shift system functions easily between issuers and acquirers, and thus to change  $C^i$  and  $C^a$  at will by the same absolute amounts but in opposite directions, there would be no need for a separate interchange fee. As Rochet and Tirole [2000] stress, however, some important functions — such as dealing with consumer default or merchant-based fraud — are more efficiently handled by one side of the system or the other. Accordingly,  $C^i$  and  $C^a$  are treated as fixed.

System behavior is modeled throughout as a two-stage game. In the second stage, the acquirer chooses  $P^a$  to maximize  $\Pi^a$ , treating  $T$  as fixed. Simultaneously, the issuer chooses  $P^i$  to maximize  $\Pi^i$ , treating  $T$  as fixed. When there are multiple issuers and/or acquirers, each takes  $T$  and all the others'  $P$ s as fixed. These are textbook problems, with objective functions that are concave under standard assumptions.

In the first stage, the interchange fee is chosen to maximize the system's *private value*,

---

<sup>11</sup> I ignore throughout the additional complications that may arise if issuers and acquirers participate in several payment systems. In the U.S., for instance, commercial banks provide both checks and credit cards. Moreover, through an arrangement referred to as “duality,” most U.S. banks issue both Visa and MasterCard cards. Some implications of duality are discussed in Evans and Schmalensee [1999, ch. 8].

<sup>12</sup> I have explored generalizations of this framework in which the acquirer can invest a non-negative amount,  $F^a$ , in marketing to build merchant demand, and the issuer can invest a non-negative amount,  $F^i$ , to build consumer demand. Then the two objective functions become  $\Pi^a = [Q^m(P^a, F^a)Q^c](P^a - C^a - T) - F^a$ , and  $\Pi^i = [Q^c(P^i, F^i)Q^m](P^i - C^i + T) - F^i$ . Unfortunately, analysis of the choice of marketing outlays in this framework, even if price competition is assumed away, turns out to involve a high ratio of technical difficulty to added insight. Accordingly, I assume  $F^a = F^i = 0$  in the text and confine discussion of results involving marketing spending to footnotes.

$$(5) \quad V = \alpha \Pi^a + (1-\alpha) \Pi^i, \text{ for } 0 \leq \alpha \leq 1.$$

Because of the multiplicative demand structure assumed here,  $V(T)$  is *not* generally globally concave, even for well-behaved partial demands. Except in pathological cases, however,  $V$  will be smooth, and values of  $T$  large enough in absolute value to drive either partial demand close to zero will not be optimal because total output will also be close to zero. Thus one of the solutions to the first-order condition  $dV/dT = 0$  will normally signal the global maximum.

If side payments were possible, it would be natural to set  $\alpha = 1/2$  in (5) and assume maximization of total system profit. But side payments are typically not possible;  $\alpha = 1/2$  is thus not necessarily descriptive of actual systems; and departures from this symmetric case are instructive. Because determining the interchange fee requires collective decision-making, which may be quite unwieldy and time-consuming, it is natural to model  $T$  as being set to maximize  $V$  in the game's first stage, before the individual banks' pricing decisions.

The next section examines what can be said about this model without specifying the functional forms of the partial demands and shows that the interchange fee plays a very different role from an ordinary market price. Section IV considers in depth the tractable case in which issuers' and acquirers' partial demand functions are linear. The welfare and output consequences of setting  $T$  to maximize private value are considered, as are the welfare and output implications of replacing cooperative systems with proprietary systems. Section V summarizes some implications of this analysis.

### III. GENERAL DEMANDS

#### *Double Marginalization*

When there is market power on both issuing and acquiring sides of the system, as we generally assume here, there is always a form of double marginalization (or, somewhat more precisely, uncoordinated pricing of complements) in the second stage of the game described above.<sup>13</sup> The interchange fee cannot help with this problem.

To see the double marginalization problem, consider the impact on  $V$  of a change in  $P^a$  at equilibrium:

$$(6) \quad \frac{\partial V}{\partial P^a} = \alpha \frac{\partial \Pi^a}{\partial P^a} + (1 - \alpha) \frac{\partial \Pi^i}{\partial P^a} = (1 - \alpha)(P^i - C^i + T) Q^c \frac{\partial Q^m}{\partial P^a}.$$

The second equation holds because  $\partial \Pi^a / \partial P^a = 0$  at equilibrium. The right-hand side of this equation will be negative as long as the issuer's margin is positive, a condition that must hold in a sustainable equilibrium with issuer non-zero market power. Since the analysis is symmetric, it follows that *at a bilateral monopoly equilibrium, small reductions in  $P^a$  and/or  $P^i$  would increase  $V$ .*<sup>14</sup>

The interchange fee cannot contribute to the solution of this double marginalization problem. Because it can only shift costs from one side of the system to the other, the interchange fee can only mitigate problems caused by *differences* between the issuing and acquiring sides. The obvious ways to deal with the double marginalization problem are to build

<sup>13</sup> This is a special case of the moral hazard problem analyzed by Holmstrom [1982].

<sup>14</sup> When marketing is possible, as discussed in note 12 above, an exactly parallel analysis demonstrates that *at a bilateral monopoly equilibrium, small increases in  $F^a$  and/or  $F^i$  would also increase  $V$ .*

a unitary, proprietary system (like American Express or Discover in the U.S.) or to have competition in both issuing and acquiring (like Visa and MasterCard in the U.S.).<sup>15</sup>

### *Bilateral Monopoly*

When  $\alpha = 1/2$ , the first-stage bilateral monopoly objective function can be written as

$$(7) \quad V = \{Q^m [P^a(T)] Q^c [P^i(T)]\} \{P^a(T) + P^i(T) - C^a - C^i\}.$$

If  $dP^a/dT = -dP^i/dT$ , shifting unit cost from one side of the market to the other leaves the second term on the right of (7) unchanged, and maximization of the first term,  $Q^T$ , is necessary and sufficient for maximization of total profit. Since  $dP^a/dT = -dP^i/dT = 1/2$  when demands are linear, it follows that *under bilateral monopoly and linear demand, the interchange fee that maximizes total profit also maximizes total system output.* (Section IV shows that Marshallian welfare,  $W$ , is also maximized in this case.)

When changes in  $T$  do affect the second term in (7), profit maximization does not imply output maximization. The difference between profit maximization and output maximization depends on exactly how the partial demand functions depart from linearity, and the profit-maximizing  $T$  may be above or below the output-maximizing fee. Similarly, the differences between these two quantities and the interchange fee that maximizes Marshallian welfare depend in general on the details of the partial demand functions.

---

<sup>15</sup> A simple example may help fix ideas. Suppose that  $C^a = C^i = 0$ ,  $Q^m = 1 - P^a$ , and  $Q^c = 1 - P^i$ . Then with  $\alpha = 1/2$ ,  $V$  is maximized by setting both  $P^a = P^i = 1/3$ , while, because of double marginalization, the equilibrium in the two-stage game has  $T = 0$  and  $P^a = P^i = 1/2$ . Now suppose that there are two issuers, 1 and 2, facing demands given by  $q_h = Q^m[(1/2) - (1/2)P_h + \gamma(P_j - P_h)]$ , with  $\gamma > 0$ , for  $h, j = 1, 2$ . Note that when the issuers' prices are equal, total demand ( $q_1 + q_2$ ) is the same as in the monopoly case. (A generalization of this setup is analyzed in Section IV.) With zero cost, Bertrand equilibrium involves  $P^a = P^i = 1/[2(1 + \gamma)]$ . If  $\gamma = 1/2$  on both sides of the system, this implies  $P^a = P^i = 1/3$ , and total profit is maximized. Output is always higher under bilateral duopoly than under bilateral monopoly in this example, and for  $\gamma < 1.618$  total profits are higher as well.

Continuing to analyze the first-stage choice of  $T$ , substitute equations (4) into equation (5), assume bank-level profit maximization at the second stage, and differentiate totally with respect to  $T$ :

$$(8) \quad \frac{dV}{dT} = \alpha \left[ \frac{\partial \Pi^a}{\partial T} + \frac{\partial \Pi^a}{\partial Q^c} \frac{dQ^c}{dT} \right] + (1 - \alpha) \left[ \frac{\partial \Pi^i}{\partial T} + \frac{\partial \Pi^i}{\partial Q^m} \frac{dQ^m}{dT} \right], \text{ or}$$

$$\frac{dV}{dT} = (1 - 2\alpha)Q^T + \alpha(P^a - C^a - T)Q^m \frac{dQ^c}{dT} + (1 - \alpha)(P^i - C^i + T)Q^c \frac{dQ^m}{dT},$$

where  $\partial \Pi^a / \partial P^a = \partial \Pi^i / \partial P^i = 0$  under bilateral monopoly by second-stage profit-maximization.

The first term on the right of this equation illustrates that if total profit is not maximized, one role played by a positive interchange fee is that of a tax levied by issuers on acquirers. The more weight the acquirer has in the system's objective function under bilateral monopoly, the less value attaches to the revenue transfer to the issuer that this tax accomplishes.

When  $\alpha = 1/2$ , so total profit is being maximized, equation (8) becomes

$$(9) \quad (P^a - C^a - T)Q^m \frac{dQ^c}{dT} + (P^i - C^i + T)Q^c \frac{dQ^m}{dT} = 0$$

Except in pathological cases, routine comparative statics analysis establishes that  $dQ^c/dT > 0$  and  $dQ^m/dT < 0$  under bilateral monopoly.<sup>16</sup> Now suppose that the two demand functions are identical. Then setting  $T = (C^a + C^i)/2$  so that unit costs are equalized ensures that  $P^a = P^i$  and that equation (7) is satisfied. Thus *under bilateral monopoly, when demand functions are*

<sup>16</sup> This statement is true in the alternative linear demand structure analyzed in Appendix B, as shown there, as long as the spillover effect coefficients introduced there are not too large. When marketing is possible, as discussed in note 12, above, equation (9) still holds, but it is much harder in this case to sign these two derivatives. It is easy to use a revealed preference argument to show that if, say,  $Q^m$  is held constant and  $T$  is increased, so the issuer's effective unit cost is reduced, the monopoly issuer will change  $P^i$  and  $F^i$  so as to increase  $Q^c$ . Similarly, if  $Q^c$  is held constant, an increase in  $T$  will decrease  $Q^m$ . But a decrease in  $Q^m$  lowers  $\partial \Pi^i / \partial F^i$ , thus lowering the optimal  $F^i$  and tending to lower  $Q^c$ . Similarly, an increase in  $Q^c$  makes acquirer marketing more attractive and thus tends to raise  $Q^m$ . There are, no doubt, stability conditions ensuring that, despite these feedbacks, an

identical, regardless of the level of (any measure of) collective market power, the necessary condition for profit maximization is satisfied when  $T$  is set to equalize issuer and acquirer unit costs. (Section IV shows that this is sufficient when partial demands are linear.)

#### *Other Market Structures*

When there are multiple issuers and acquirers, the expression for  $dV/dT$  is in general more complex than (8) because neither side of the system maximizes its total profit in the game's second stage. Moreover, there is no completely general guarantee that changes in the interchange fee will raise one partial demand and lower the other. Following Dixit [1986], however, the Appendix demonstrates that in a substantial class of Bertrand oligopoly models, increases in unit cost lower total market demand when standard stability conditions are imposed.<sup>17</sup> Thus the choice of the interchange fee under oligopoly or monopolistic competition generally involves a tradeoff between the partial demands of issuers and acquirers.

#### IV. LINEAR DEMANDS

This section considers the tractable case of linear partial demand functions. Suppose there are  $N$  firms on one side of the system, with demands given by

$$(10) \quad q_i = Q^{oe} \left\{ \frac{A}{N} + \Theta \left[ \sum_{j \neq i} P_j - (N-1)P_i \right] - \frac{B}{N} P_i \right\}, \quad i = 1, \dots, N,$$

where  $Q^{oe}$  is expected partial demand on the other side of the system. (Superscripts are dropped in most of this paragraph and the next to avoid clutter.) The sum of the  $q_i$  will equal

---

increase in  $T$  will raise  $Q^c$  and lower  $Q^m$  in equilibrium when marketing is possible, but I have not attempted to derive those conditions.

<sup>17</sup> It is easy to show that this is also true in the marketing competition model (with fixed prices) of Schmalensee [1976], when the stability conditions derived there are assumed.

$Q^{oe} [A - B\bar{P}]$ , where  $\bar{P}$  is the average of the  $P_i$ . The larger is  $\Theta$ , the more sensitive market shares are to differences in prices. Appendix B considers a generalization of this system, in which the expression in brackets in (10) depends directly on  $Q^{oe}$ . This generalization allows for more complex patterns of network effects in fulfilled expectations equilibrium, because  $Q^{oe}$  affects second-stage pricing, and adds considerable algebraic complexity, but it does not change the basic economics of the system.

Suppose all firms on this side of the market have unit cost  $C'$ , net of interchange. (If this is the acquiring side of the system,  $C' = C^a + T$ , while  $C' = C^i - T$  on the issuing side.)

Multiplying (10) by  $(P_i - C')$ , differentiating, and solving for a symmetric equilibrium yields

$$(11) \quad Q = \frac{(A - BC')(B + N(N-1)\Theta)}{2B + N(N-1)\Theta},$$

where  $Q$  is the total partial demand on this side of the system, the sum of the  $q_i/Q^{oe}$ , and

$$(12) \quad P - C' = \frac{(A - BC')}{2B + N(N-1)\Theta}.$$

Now suppose that there are  $N^a$  acquirers, with linear demands as above and net unit costs equal to  $(C^a + T)$ , and  $N^i$  issuers, with linear demands as above and net unit costs equal to  $(C^i - T)$ . The results of the preceding paragraph imply that at a symmetric equilibrium,

$$(13a) \quad Q^m = \frac{(D^m - B^m T)(B^m + \beta^m)}{2B^m + \beta^m}, \quad Q^c = \frac{(D^c + B^c T)(B^c + \beta^c)}{2B^c + \beta^c}; \text{ and}$$

$$(13b) \quad P^a - C^a - T = \frac{D^m - B^m T}{2B^m + \beta^m}, \quad P^i - C^i + T = \frac{D^c + B^c T}{2B^c + \beta^c}; \text{ where}$$

$$(13c) \quad \beta^m = N^a(N^a - 1)\Theta^m, \quad \beta^c = N^i(N^i - 1)\Theta^c,$$

$$(13d) \quad D^m = A^m - B^m C^a, \text{ and } D^c = A^c - B^c C^i.$$



The larger is  $\beta^m$  ( $\beta^c$ ), the more intense is competition among acquirers (issuers).

### *Output and Welfare*

From (13a), total system output is a quadratic in  $T$ , which is maximized at

$$(14) \quad T^Q \equiv \frac{1}{2} \left( \frac{D^m}{B^m} - \frac{D^c}{B^c} \right) = \frac{1}{2} \left[ \left( \frac{A^m}{B^m} - \frac{A^c}{B^c} \right) + (C^i - C^a) \right].$$

Note first that *if partial demand functions are linear and identical, it is output-maximizing to choose  $T$  to equalize unit costs*. More generally, all else equal, it is output-maximizing for interchange to flow to the high-cost side of the system, which would otherwise find it more difficult to stimulate total system demand. It is easy to show that the higher is  $(A^h/B^h)$ , for  $h=m,c$ , the lower the elasticity of  $Q^h$  with respect to  $P^h$  at any given  $P^h$ . Thus equation (14) implies that *the more elastic the issuers' demand is relative to the acquirers' demand, the higher the output-maximizing interchange fee* (which is paid to the issuers).<sup>18</sup> The intuition is that it is output-maximizing to subsidize price cuts where they will do the most good to increase output for the system as a whole, and that is where demand is more elastic. Unless the partial demand functions are identical, using cost-based regulation to determine  $T$  will maximize system output only by chance.

To analyze Marshallian social welfare,  $W$ , defined by equation (3), it is first necessary to invert the demand system (8) and integrate to obtain the corresponding partial equilibrium surplus function. Confining attention to symmetric equilibria, at which the  $q_i$  on each side of the system are equal and using (13a), we obtain

---

<sup>18</sup> The conclusion that only elasticity differences matter flows from the assumption that the functional forms of the two partial demand functions are the same. In general, differences in functional forms will also affect the impact on total demand of changes in the interchange fee.

$$\begin{aligned}
(15) \quad W &= \left[ \frac{A^M}{B^m} Q^T - \frac{1}{2B^m Q^c} (Q^T)^2 \right] + \left[ \frac{A^c}{B^c} Q^T - \frac{1}{2B^c Q^m} (Q^T)^2 \right] - (C^a + C^i) Q^T \\
&= \frac{Q^T}{2} \left\{ \left[ \frac{D^m (3B^m + \beta^m)}{B^m (2B^m + \beta^m)} + \frac{D^c (3B^c + \beta^c)}{B^c (2B^c + \beta^c)} \right] + \left[ \frac{B^m + \beta^m}{2B^m + \beta^m} - \frac{B^c + \beta^c}{2B^c + \beta^c} \right] T \right\}, \\
&= \frac{Q^T}{2} (\delta + \lambda T),
\end{aligned}$$

where the final equality defines  $\delta$  and  $\lambda$ . Note that  $\delta > 0$ , while  $\lambda$  has the sign of  $[(\beta^m/B^m) - (\beta^c/B^c)]$ . When there is only one issuer and one acquirer,  $\lambda = 0$ , so that *under bilateral monopoly maximizing system output is equivalent to maximizing Marshallian social welfare*. More generally,  $dW/dT$  is a quadratic in  $T$  that with roots that resist simplification. When  $Q^T$  is maximized, however,  $dW/dT$  has the sign of  $\lambda$ , and it follows easily that if  $T^W$  maximizes  $W(T)$ ,

$$(16) \quad [T^W - T^Q][(\beta^m/B^m) - (\beta^c/B^c)] \geq 0.$$

The discussion below shows that the difference between  $T^W$  and  $T^Q$  reflects that fact that Marshallian welfare depends on profit as well as consumers' surplus.

### *Private Value*

Substituting equations (13) into equations (4) and (5), it is easy to show that private value,  $V$ , is proportional to

$$(17a) \quad V' = \frac{\omega}{B^m} (D^c + B^c T)(D^m - B^m T)^2 + \frac{1-\omega}{B^c} (D^m - B^m T)(D^c + B^c T)^2, \text{ where}$$

$$(17b) \quad \omega = \alpha \frac{2 + (\beta^c / B^c)}{2 + \alpha(\beta^c / B^c) + (1 - \alpha)(\beta^m / B^m)}.$$

Note that  $\omega = \alpha$  under bilateral monopoly, when  $\beta^m = \beta^c = 0$ . More generally, the larger is  $(\beta^c/B^c)$  or the smaller is  $(\beta^m/B^m)$ , the larger is  $\omega$ . Differentiating (17a) yields the first-order necessary condition for first-stage maximization of private value:

$$(18) \quad \frac{dV'}{dT} = 2(1 - 2\omega)(D^m - B^m T)(D^c + B^c T) + \frac{\omega B^c}{B^m} (D^m - B^m T)^2 - \frac{(1 - \omega)B^m}{B^c} (D^c + B^c T)^2 = 0$$

When  $\omega = 1/2$ , equation (18) has a single real root,  $T^V(1/2)$ , that corresponds to a maximum of system value, and  $T^V(1/2) = T^Q$ , where  $T^Q$  is defined by equation (14). That is, *profit maximization under bilateral monopoly, or, more generally, private value maximization with  $\omega=1/2$ , implies maximization of total system output. Under bilateral monopoly, Marshallian social welfare is also maximized.* As in Section III, the intuition is that increasing total output, by moving units costs toward equality and subsidizing price cuts where demand elasticity is high, increases the size of the pie for the system as a whole.<sup>19</sup>

Moreover, equation (14) shows that *the privately optimal interchange fee when  $\omega=1/2$  depends only on differences between the two sides of the system, not on any measure of the level of market power.* If costs and partial demand functions are identical, for instance, the optimal interchange fee is zero no matter how much or how little market power the system as a whole enjoys. Alternatively, under bilateral monopoly with  $B^m = B^c$ , it is easy to show that the maximum level of system profit, a plausible measure of market power, varies with  $(D^m + D^c)$ , while the profit-maximizing interchange fee,  $T^V(1/2) = T^Q$ , varies with  $(D^m - D^c)$ .

When  $\omega \neq 1/2$ , equation (18) has two real roots. The root corresponding to a maximum of  $V$  is

---

<sup>19</sup> Following the discussion in note 12, above, I have investigated a bilateral monopoly model in which prices are fixed, the issuer and acquirer choose fixed costs, and partial demands are given by  $\ln(Q^m) = \phi^m \ln(F^a)$ , and  $\ln(Q^c) = \phi^c \ln(F^i)$ , with  $\phi^m$  and  $\phi^c$  constants between zero and one. (The basic structure comes from Schmalensee [1976].) Numerical experiments suggest that in this model profit-maximizing interchange tends to flow, all else equal, to the side of the system with the smallest price-cost margin and to the side for which demand is more sensitive to fixed cost outlays (i.e., the side with the larger value of  $\phi$ ).

$$(19) \quad T^V(\omega) = T^Q + \left( \frac{D^m}{B^m} + \frac{D^c}{B^c} \right) \frac{1 - \sqrt{1 + 12(\omega - 1/2)^2}}{12(\omega - 1/2)}.$$

Since  $T^V(\omega)$  is a decreasing function, from (17b) *the private value-maximizing interchange fee is a decreasing function of  $\alpha$  and  $(\beta^c/B^c)$ , and an increasing function of  $(\beta^m/B^m)$ .*

Under profit maximization, when  $\alpha = 1/2$ ,  $(\omega - 1/2)$  has the sign of  $[(\beta^c/B^c) - (\beta^m/B^m)]$ .

Comparing (16), if  $T^\Pi$  is the value of  $T$  that maximizes total system profit,

$$(20) \quad [T^\Pi - T^Q][T^W - T^Q] \geq 0$$

That is, *the profit-maximizing  $T$  departs from  $T^Q$  in the same direction as the welfare-maximizing  $T$ .* This result, which echoes the relation between Ramsey pricing and monopoly price discrimination and similarly reflects the inclusion of profits in the Marshallian welfare measure, does not seem to be easily generalized beyond the linear case. From (13b), with linear partial demands the more intense is competition on either side of the system, the less sensitive is the unit markup on that side of the system to changes in  $T$ . (In the limit as  $\beta$  increases, unit markup goes to zero, independent of  $T$ .) It is easy to show that the derivative of total system markup,  $[(P^a + P^i) - (C^a + C^i)]$ , with respect to  $T$  has the sign of  $[(\beta^m/B^m) - (\beta^c/B^c)]$ . Thus if  $B^m = B^c$ , for instance, and there is more intense competition on the issuing side than on the acquiring side ( $\beta^m < \beta^c$ ), it is both profit-maximizing and welfare-maximizing to reduce  $T$  below the output-maximizing level in order to increase total system markup.

When  $\alpha \neq 1/2$ , the interchange fee is affected by the desirability of shifting profit from one side of the system to the other. Under bilateral monopoly, with  $\omega = \alpha$ , the second term on the right of (19) directly reflects the use of the interchange fee to transfer profit from one side of the system to the other. When  $\alpha < 1/2$ , for instance, so that the issuer's profit is weighted

more heavily than the acquirer's, this second term is positive. In this case  $T$  is increased, all else equal, in order to transfer profit to the issuer, and, all else equal, system output and welfare are reduced as a consequence.

### *Alternative System Structures*

In the U.S., banks' voting power in the Visa and MasterCard associations is more sensitive to issuing volume than to acquiring volume, indicating  $\alpha < 1/2$ . In addition, the acquiring side of the U.S. bank credit card business involves little or no product differentiation and is generally viewed as highly competitive, indicating  $\beta^m$  is large.<sup>20</sup> From equation (17b), this suggests that the polar case  $\omega = 0$  is of particular interest.<sup>21</sup> In this case, equations (14) and (19) directly imply

$$(21) \quad T^V(0) = \frac{1}{3} \left[ \left( 2 \frac{A^m}{B^m} - \frac{A^c}{B^c} \right) + (C^i - 2C^a) \right] > T^Q.$$

Note that  $T^V(0)$  is independent of  $\beta^c$  and thus of the intensity of competition among issuers.

Even though in this polar case acquirers' cost and demand conditions are weighted more heavily than those of issuers, differences between the two sides of the system remain central, and the qualitative impacts of changes in cost and demand conditions are essentially the same as under output maximization.

To evaluate the importance of this extreme departure from output maximization in a cooperative system, let  $Q^{MAX}$  be the maximum value of total system output:

---

<sup>20</sup> Structurally, the acquiring business does not look perfectly competitive. (See, generally, Evans and Schmalensee [1999, ch. 6].) Most U.S. banks contract out this function to specialists, and, because of scale economies in transaction processing, concentration in acquiring is relatively high. Still, competition in this commodity business is generally described as intense, and margins are small relative to, e.g., interchange fees, so that perfect competition may be a good behavioral approximation.

$$(22a) \quad Q^{MAX} = \left[ \frac{(B^m + \beta^m)(B^c + \beta^c)}{(2B^m + \beta^m)(2B^c + \beta^c)} \right] \left[ \frac{(D^m B^c + D^c B^m)^2}{4B^m B^c} \right] = K \left[ \frac{(D^m B^c + D^c B^m)^2}{4B^m B^c} \right],$$

where the second equality defines  $K$ . In general  $K$  is between 1/4 and one, depending on competitive conditions among issuers and acquirers. When  $\beta^n$  is large, as was assumed in deriving (21),  $K$  is between 1/2 and one.

Substitution of (21) into equations (13a) yields total system output when  $T = T^V(0)$ :

$$(23) \quad Q^V(0) = (8/9)Q^{MAX}.$$

That is, *when the interchange fee is at the highest value consistent with private value maximization, total output is reduced by about 11 percent from its maximum value.*

To put this reduction in perspective and to shed light on some current controversies, it is useful to consider total output under alternative system structures. Consider first a non-unitary proprietary system, which charges acquirers a fee  $T^a$  per transaction, charges issuers a fee  $T^i$  per transaction, and sets these fees to maximize  $(T^a + T^i)Q^T$ . It is straightforward to show that the corresponding total output level is given by

$$(24) \quad Q^{PN} = (4/9)Q^{MAX} = (1/2)Q^V(0).$$

That is, *moving from a cooperative system to a non-unitary proprietary system, keeping the numbers of (independent) issuers and acquirers constant, reduces output by between 50 and 56 percent.* This result should make clear the fundamental economic difference between an interchange fee passed from acquirers to issuers in a cooperative system and an ordinary per-transaction fee set by a proprietary system to maximize its profit.

---

<sup>21</sup> As noted above, this is in effect the case on which Rochet and Tirole [2000] focus.

Finally, consider a unitary proprietary system, which does all its own issuing and acquiring and sets  $P^a$  and  $P^j$  to maximize total system profit. The corresponding total output level is given by

$$(25) \quad Q^{PI} = (1/K)Q^{PN} = (4/9K)Q^{MAX} = (1/2K)Q^V(0).$$

*Except in the case of perfect competition in issuing and acquiring (when  $K = 1$ ), total output for a unitary proprietary system exceeds that for a non-unitary proprietary system, all else equal.* The non-unitary system's profit is, in effect, the receipts from taxing an imperfectly competitive market, thus giving rise to a double marginalization problem. On the other hand, *if competition in issuing and/or acquiring is vigorous, so  $K > 1/2$ , a unitary proprietary system always has lower total output than a cooperative system, all else equal.*

## V. IMPLICATIONS

The policy question motivating this paper is whether antitrust authorities should condemn collective determination of interchange fees for the same reasons they would condemn competing banks fixing credit card interest rates or annual fees.<sup>22</sup> The analysis here provides no support for such a policy. The interchange fee is not an ordinary market price; it is a balancing device for increasing the value of a payment system by shifting costs between issuers and acquirers and thus shifting charges between consumers and merchants.<sup>23</sup> The first-order effect of fixing an ordinary price is to harm consumers by reducing output, while in a non-

<sup>22</sup> As discussed in note 3, above, the formal analysis here does not deal with the argument that merchant discounts should be reduced (at least in part by putting pressure on interchange fees) in order to reduce distortions in retail pricing.

<sup>23</sup> Balto [2000] and others who condemn collective determination of interchange fees seem to ignore this balancing role. Thus they condemn interchange fee increases because they raise merchants' costs and forget that the same logic says that interchange fee increases lower consumers' costs.

extreme case collective interchange fee determination *maximizes* output and Marshallian welfare in order to maximize the system's private value to its owners.

More generally, our analysis shows that both the private value-maximizing interchange fee and the output-maximizing fee are determined mainly by *differences* between issuers and acquirers; symmetry makes a zero interchange fee optimal. The model employed here is thus consistent with collectively determined interchange flowing to either issuers or acquirers, and we observe both patterns in reality. We find that the private value-maximizing interchange fee may be above or below the output-maximizing fee and that the welfare-maximizing fee differs from the output-maximizing fee in the same direction as the profit-maximizing fee does. Increasing the interchange fee from its privately optimal level may increase total system output and, and decreasing it may decrease output.

Even in the special case of linear partial demands, our analysis reveals no straightforward policy toward the interchange fee that can be reliably expected to improve system performance on balance.<sup>24</sup> Small and Wright [2000] have argued that moving interchange fees from collective determination to bilateral negotiations could raise fees on average, with unpredictable impacts on output and welfare. Similarly, if interchange fees were set to zero, nothing in this analysis suggests that total system output or welfare would be more likely to rise than to fall. Except in *very* special circumstances, no cost-based approach to regulating interchange fees can guarantee to increase system output or Marshallian welfare over

---

<sup>24</sup> As discussed in note 3, above, some observers contend that because retailers partially shift merchant discounts to consumers using cash and checks, value-maximizing credit card systems may set interchange fees inefficiently high. Even if this argument were generally correct, despite the second-best issues raised in note 3 and the complexities discussed in the text, substantial reductions in interchange fees may well reduce card system output substantially, directly harming consumers. Thus it does not follow that reducing interchange fees to zero (or some cost-driven level) can be expected to make consumers better off on balance.



private value-maximizing levels. It is highly unlikely that regulators would ever have enough information to implement the socially optimal interchange fees discussed in Section IV and Appendix B, and these solutions rest on a set of restrictive assumptions.

Despite these uncertainties, any serious restriction on collective interchange fee determination would have one clear effect: it would make it harder for the bank card systems to compete effectively with American Express and other proprietary payment systems. As I noted above, because within the U.S. it does all its system's issuing and acquiring, American Express has been free to set merchant discounts and cardholder fees there without fear of antitrust attack. It has generally chosen to set merchant discounts that could be matched by the bank card systems only if they were to *raise* their interchange fees substantially.<sup>25</sup> In some other countries it has operated as a non-unitary proprietary system. Because the fees specified in its contracts with independent issuers and acquirers were not the result of agreements between competitors, however, they have also been immune to antitrust attack.

Barring collective interchange fee determination would create strong incentives for large institutions to abandon the cooperative bank card systems and create proprietary systems. As the analysis of Section IV indicated, however, all else equal a movement from cooperative to proprietary systems is likely to reduce total system output. All in all, there is no economic defense for an antitrust policy favoring proprietary payment systems over cooperative payment systems pursuing broadly similar strategies.

---

<sup>25</sup> Discover/Novus has charged a lower average merchant discount than Visa or MasterCard acquirers, though its average merchant discounts have exceeded the markups charged by bank card system acquirers over the bank card systems' interchange fees and thus likely have exceeded the levels that would emerge in the bank card systems of interchange fees were forced to zero.

## APPENDIX A

Consider a market with  $N$  firms selling differentiated products in which firm  $i$ 's demand,  $q_i$ , depends only on its price,  $p_i$ , and the *average* price of its  $N-1$  rivals,  $p_{-i}$ . (As Dixit [1986, p. 119] notes, without some restrictive assumption of this sort, it is generally not possible to do comparative statics in differentiated product oligopolies.) If  $c_i$  is firm  $i$ 's constant unit cost, the set of first-order conditions that must be satisfied at a Bertrand equilibrium is

$$(A1) \quad \mu_i = q_i + (p_i - c_i) \sigma_i = 0, \quad i = 1, \dots, N,$$

where  $\sigma_i \equiv \partial q_i / \partial p_i < 0$ ,  $i = 1, \dots, N$ . Key quantities in the analysis that follows are

$$(A2) \quad a_i \equiv \partial \mu_i / \partial p_i < 0, \quad i = 1, \dots, N; \text{ and}$$

$$(A3) \quad b_i = \frac{\partial \mu_i}{\partial p_j} = \frac{\sigma_{-i} + (p_i - c_i) \frac{\partial \sigma_i}{\partial p_{-i}}}{N-1}, \quad i = 1, \dots, N, \quad j \neq i,$$

where  $\sigma_{-i} \equiv \partial q_i / \partial p_{-i}$ ,  $i = 1, \dots, N$ . The sign of the  $a_i$  follows from the second-order conditions.

The natural assumption that competing products are (gross) substitutes implies that  $\sigma_{-i} > 0$ , but there is no obvious reason why  $\partial \sigma_i / \partial p_{-i} = \partial \sigma_{-i} / \partial p_i$  should be positive or negative. I assume that the  $\sigma_{-i}$  terms dominate, so that  $b_i > 0$  for all  $i$ .

In order to do comparative statics in oligopoly models, it is generally necessary to invoke stability conditions to replace the ‘‘off-diagonal’’ second order conditions that arise in monopoly models. (See Schmalensee [1976], Dixit [1986], and the references they cite.) Here, following Dixit [1986, p. 117], I assume the following diagonal dominance condition is satisfied:

$$(A4) \quad a_i + (N-1)b_i < 0, \quad i = 1, \dots, N.$$

This is a sufficient condition for stability under conventional dynamic assumptions.

Now suppose that  $c_i$  is replaced throughout by  $(c_i + \theta)$ . The goal here is to sign  $dQ/d\theta$  at  $\theta=0$ , where  $Q$  is the sum of the  $q_i$ . I do this by showing that a small increase in  $\theta$  raises all prices in equilibrium. Letting “ $dx$ ” be shorthand for “ $dx/d\theta$  at  $\theta=0$ ,” totally differentiate the first-order conditions (A1) to obtain

$$(A5) \quad a_i(dp_i) + [(N-1)b_i](dp_{-i}) - \sigma_i = 0, \quad i = 1, \dots, N.$$

Without loss of generality, suppose  $dp_1 \leq dp_2 \leq \dots \leq dp_N$ . This implies that  $dp_{-1} \geq dp_{-2} \geq \dots \geq dp_{-N}$ .

To show that all the  $dp_i$  are positive under the above assumptions, let us suppose  $dp_1 \leq 0$  and show a contradiction. Since  $a_1 < 0$ ,  $b_1 > 0$ , and  $\sigma_1 < 0$ , equation (A5) shows that  $dp_1 \leq 0$  implies  $dp_{-1} < 0$ . It then follows from the inequalities just above that all other  $dp_{-i}$  must also be negative.

Dividing (A5) by  $a_i$  and summing across all firms in the market yields

$$(A6) \quad \sum_{i=1}^N dp_{-i} \left[ 1 + \frac{(N-1)b_i}{a_i} \right] = \sum_{i=1}^N \frac{\sigma_i}{a_i}.$$

The summation on the right is positive, and so, from (A4), are each of the terms in brackets on the left. It is accordingly not possible for (A6) to hold if all the  $dp_{-i}$  are all negative. The assumption that one of the prices does not increase has thus led to a contradiction, so all prices must rise when unit costs increase across the board, and total output must accordingly fall.

## APPENDIX B

This appendix generalizes the linear demand system analyzed in Section IV to allow partial demands to exhibit network effects directly, so that the state of the other side of the system affects pricing, and analyzes the relation between maximizing output and maximizing private

value. Consider one side of the system with  $N$  firms and with demands given by a simple generalization of (10):

$$(B1) \quad q_i = Q^{oe} \left\{ \frac{A}{N} + \frac{\varphi Q^{oe}}{N} + \Theta \left[ \sum_{j \neq i} P_j - (N-1)P_i \right] - \frac{B}{N} P_i \right\}, \quad i = 1, \dots, N,$$

where  $Q^{oe}$  is the expected partial demand on the other side of the system, and  $\varphi$  is a positive constant. (Superscripts are generally dropped in this paragraph to avoid clutter.) The larger is  $\varphi$ , the more important are direct network effects on this side of the system. Suppose all firms on this side of the system have unit cost  $C'$ , net of interchange. (If this is the acquiring side of the system,  $C' = C + T$ , while  $C' = C - T$  on the issuing side.) Multiplying (B1) by  $(P_i - C')$ , differentiating, and solving for a symmetric equilibrium conditional on  $Q^{oe}$ , we obtain

$$(B2) \quad Q = \frac{(A + \varphi Q^{oe} - BC')(B + \beta)}{(2B + \beta)},$$

where, as in the text,  $\beta = N(N-1)\Theta$ .

To obtain a symmetric fulfilled expectations equilibrium, one must solve for conditional equilibrium on both sides of the system, as just above, set expected partial demands equal to actual partial demands, and solve the resulting pair of linear equations. In order for this solution to involve positive partial demands, network effects must not be too large:

$$(B3) \quad \varphi^m \varphi^c < \frac{(2B^m + \beta^m)(2B^c + \beta^c)}{(B^m + \beta^m)(B^c + \beta^c)}.$$

Prices are obtained by inverting the two partial demand functions, substituting actual for expected partial demands.

If there are  $N^a$  acquirers, with linear demands as above and net unit costs equal to  $(C^a+T)$ , and  $N^i$  issuers, with linear demands as above and net unit costs equal to  $(C^i-T)$ , the fulfilled expectations equilibrium is as follows:

$$(B4a) \quad Q^m = \frac{(d^m - b^m T)(B^m + \beta^m)}{Z}, Q^c = \frac{(d^c + b^c T)(B^c + \beta^c)}{Z}; \text{ and}$$

$$(B4b) \quad P^a - C^a - T = \frac{d^m - b^m T}{Z}, P^i - C^i + T = \frac{d^c + b^c T}{Z};$$

where  $\beta^m, \beta^c, D^m$ , and  $D^c$  are defined by equations (13c) and (13d), above,

$$(B4c) \quad d^m = D^m(2B^c + \beta^c) + \phi^m D^c(B^c + \beta^c), d^c = D^c(2B^m + \beta^m) + \phi^c D^m(B^m + \beta^m),$$

$$(B4d) \quad b^m = B^m(2B^c + \beta^c) - \phi^m B^c(B^c + \beta^c), b^c = B^c(2B^m + \beta^m) - \phi^c B^m(B^m + \beta^m), \text{ and}$$

$$(B4e) \quad Z = (2B^m + \beta^m)(2B^c + \beta^c) - \phi^m \phi^c (B^m + \beta^m)(B^c + \beta^c).$$

Note that condition (B3) is equivalent to  $Z > 0$ . The necessary and sufficient conditions for equilibrium  $Q^c$  to be increasing in  $T$  (i.e.,  $b^c > 0$ ) and  $Q^m$  to be decreasing in  $T$  (i.e.,  $b^m > 0$ ) also require that network effects not be too large. These two conditions, which are sufficient for  $Z > 0$ , will both be satisfied for all non-negative  $\beta^m$  and  $\beta^c$  if and only if

$$(B5) \quad \phi^c < B^c/B^m \text{ and } \phi^m < B^m/B^c.$$

As in Section IV, total system output is a quadratic in  $T$ ; it is maximized at

$$(B6) \quad T^{QN} = \frac{1}{2} \left( \frac{d^m}{b^m} - \frac{d^c}{b^c} \right).$$

This, of course, reduces to (14) when  $\phi^i = \phi^c = 0$ . Here, from (B4c) and (B4d), the output-maximizing interchange fee depends on all the cost and demand parameters in the system, including the  $\phi$ s and the  $\beta$ s.

It is somewhat surprising that the bounds on  $\varphi^m$  and  $\varphi^c$  introduced above do not generally suffice to sign the derivatives of  $T^{QN}$  with respect to  $D^m$ ,  $D^c$ ,  $B^m$ , or  $B^c$ .<sup>26</sup> On the other hand, it is easy to show that  $\partial T^{QN}/\partial \varphi^m > 0$  and  $\partial T^{QN}/\partial \varphi^c < 0$ . The larger is  $\varphi^m$ , for instance, the greater is the spillover benefit to total output from increasing  $Q^c$  by raising  $T$ , and this benefit shows up as an increase in  $Q^m$ . It is straightforward to show that  $\partial T^{QN}/\partial \beta^m < 0$  and  $\partial T^{QN}/\partial \beta^c > 0$ . All else equal, the more intense is competition on the acquiring side, for instance, the greater the reduction in  $P^m$  induced by lowering  $T$ ; and in this model a reduction in  $P^m$  increases  $Q^c$  as well as  $Q^m$ .

From equations (B4), the system's objective function,  $V$ , is proportional to

$$(B7a) \quad V'' = \frac{\omega'}{b^m} (d^c + b^c T)(d^m - b^m T)^2 + \frac{1 - \omega'}{b^c} (d^m - b^m T)(d^c + b^c T)^2, \text{ where}$$

$$(B7b) \quad \omega' = \alpha \frac{2 + [(\beta^c / B^c) - (\varphi^m / B^m)(B^c + \beta^c)]}{2 + \alpha[(\beta^c / B^c) - (\varphi^m / B^m)(B^c + \beta^c)] + (1 - \alpha)[(\beta^m / B^m) - (\varphi^c / B^c)(B^m + \beta^m)]}.$$

Comparing equations (17) and (B7), it is straightforward to characterize the privately optimal choice of interchange fee in this model. First, if  $\omega' = 1/2$ , *private value maximization implies maximization of total system output*. Second, following equation (19), *for fixed  $T^{QN}$ , the difference between the privately optimal interchange fee and the output-maximizing interchange fee is a decreasing function of  $\omega'$* . (Much as in Section IV, the derivative of total system markup,  $[(P^a + P^i) - (C^a + C^i)]$ , with respect to  $T$  has the sign of  $(b^c - b^m)$ , which is proportional to the difference between the two bracketed terms in the denominator of (B7b).)

<sup>26</sup> Under bilateral monopoly,  $T^{QN}$  is increasing in  $B^i$  and decreasing in  $B^a$ ; if (B5) is satisfied it is increasing in  $D^a$  and decreasing in  $D^i$ .

This is a relatively weak result, however, since in this model with imperfect competition on both sides of the system, both  $T^{QN}$  and  $\omega'$  depend on the  $B$ s,  $\beta$ s, and  $\varphi$ s.

The determination of  $\omega'$  via (B7b) is sufficiently complex that, almost regardless of system objectives, it would seem hard to guess whether a private value-maximizing interchange fee would be above or below the output-maximizing level.<sup>27</sup> In general,  $\omega'$  is increasing in  $\alpha$  and  $\varphi^c$  and decreasing in  $\varphi^m$ , and, if conditions (B5) are satisfied, it is increasing in  $\beta^c$  and decreasing in  $\beta^m$ . It follows that, like the output-maximizing  $T$ , the  $T$  that maximizes private value is decreasing in  $\varphi^c$  and increasing in  $\varphi^m$ .

## REFERENCES

- Balto, David A., 2000, 'The Problem of Interchange Fees: Costs Without Benefits?', *European Competition Law Review*, 21, pp. 215-224.
- Baxter, William F., 1983, 'Bank Interchange of Transactional Paper: Legal and Economic Perspectives', *Journal of Law and Economics*, 26, pp. 541-588.
- Carlton, Dennis W. and Frankel, Alan S., 1995a, 'The Antitrust Economics of Credit Card Networks', *Antitrust Law Journal*, 63, pp. 643-668.
- \_\_\_\_\_ and \_\_\_\_\_, 1995b, 'The Antitrust Economics of Credit Card Networks: Reply to Evans and Schmalensee Comment', *Antitrust Law Journal*, 63, pp. 903-915.
- Chang, Howard H., and Evans, David S., 2000, 'The Competitive Effects of the Collective Setting of Interchange Fees by Payment Card Systems', *Antitrust Bulletin*, 45, pp. 641-677.
- Competition Directorate-General, European Commission, 2000, 'Commission Plans to Clear Certain Visa Provisions, Challenge Others', Press Release IP/00/1164, October 16.
- Dixit, Avinash K., 1986, 'Comparative Statics for Oligopoly', *International Economic Review*, 27, pp. 107-122.

---

<sup>27</sup> In the polar case considered at the end of Section IV, as long as the first of conditions (B5) is satisfied,  $\beta^m \rightarrow \infty$  implies  $\omega' \rightarrow 0$ , leading to a choice of  $T$  above the output-maximizing level.

- Evans, David S., and Schmalensee, Richard L., 1993, *The Economics of the Payment Card Industry*, Cambridge, National Economic Research Associates.
- \_\_\_\_\_ and \_\_\_\_\_, 1995, 'Economic Aspects of Payment Card Systems and Antitrust Policy Toward Joint Ventures', *Antitrust Law Journal*, 63, pp. 861-901.
- \_\_\_\_\_ and \_\_\_\_\_, 1999, *Paying with Plastic*, Cambridge, MIT Press.
- \_\_\_\_\_, \_\_\_\_\_, and Chang, Howard, 1998, 'Some Economic Principles for Guiding Antitrust Policy Towards Joint Ventures', *Columbia Business Law Review*, pp. 223-329.
- Faulkner & Gray, 1999, *Debit Card Directory, 1999 Edition*, New York: Faulkner & Gray, Inc.
- Frankel, Alan S., 1998, 'Monopoly and Competition in the Supply and Exchange of Money', *Antitrust Law Journal*, 66, pp. 313-361.
- Hehir, Grainne, 2000, 'Politics & Economy: EU Launches Investigation of Visa Practices — Regulators Say Certain Fees May Violate Antitrust Rules', *Wall Street Journal Europe*, October 17.
- Holmstrom, Bengt, 1982, 'Moral Hazard in Teams', *Bell Journal of Economics*, 13, pp. 324-340.
- Jessup, P., 1967, *The Theory and Practice of Nonpar Banking*, Evanston: Northwestern University Press.
- Kim, John, 1998, 'The Impact of Proprietary Positions and Equity Interest in the Pricing of Network ATM Services', chapter in unpublished dissertation, Department of Economics, Massachusetts Institute of Technology, Cambridge.
- Reserve Bank of Australia and Australian Competition and Consumer Commission, 2000, 'Debit and Credit Card Schemes in Australia: A Study of Interchange Fees and Access,' mimeo, Canberra.
- Rochet, Jean-Charles, and Tirole, Jean, 2000, 'Cooperation Among Competitors: The Economics of Payment Card Associations', mimeo, University of Toulouse.
- Schmalensee, Richard, 1976, 'A Model of Promotional Competition in Oligopoly', *Review of Economic Studies*, 43, pp. 493-507.
- Schwartz, Marius, and Vincent, David. R., 2000. 'The No Surcharge Rule in Electronic Payments Markets: A Mitigation of Pricing Distortions?', mimeo, Georgetown University.



Small, John, and Wright, Julian, 2000, 'Decentralized Interchange Fees in Open Payment Networks: An Economic Analysis', mimeo, NECG and University of Auckland.

Spahr, W., 1926, *The Clearing and Collection of Checks*, New York: The Bankers Publishing Co.

Wright, Julian, 2000, 'An Economic Analysis of a Card Payment Network', mimeo, NECG and University of Auckland.