increases with the duration of the single relationship.'' They implied that once a firm has proved to be successful for certain number of periods, it develops multiple banking links to avoid the holdup problem. Petersen and Rajan (1994) considered the relationship between age and bank borrowing and obtained a similar result: ''The fraction of bank borrowing declines from 63 percent for firms aged 10 to 19 years to 52 percent for the oldest firms in our sample. This seems to suggest that firms follow a ''pecking order'' for borrowing over time, starting with the closest sources (family) and then progressing to more arm's length sources.'' This evidence tends to indicate that the younger the firm, the larger the degree of asymmetric information, which leaves young firms with the unique option of a single bank relationship. Ongena and Smith (2001) corroborate this finding and establish, in addition, that small, young, and highly leveraged firms maintain shorter relationships.

Finally, and directly related to the empirical evidence on the effect of distance on credit conditions (see section 3.3.4), the literature on relationship banking has addressed the issue of distance by testing whether the effects of distance on loan rates could be the result of asymmetric information. Degryse and Ongena (2005b) and Agarwal and Hauswald (2007) conclude that distance effect on loan rates decreases (or becomes irrelevant) when some measure of relationship banking is introduced into the regression. Degryse and Ongena found the loan rate charged to relationship borrowers is unaffected by distance while the rate charged to arm's-length borrowers decreases with distance. In Agarwal and Hauswald's contribution, the firm-bank distance becomes statistically insignificant when the lending bank's internal rating (the bank's proprietary information) is introduced.

In a similar vein, Berger et al. (2004) find that small banks lend at lower distances because they have a comparative advantage in gathering soft information controlling for firm and market characteristics. Berger et al. show that the distance between a bank and its borrowing firm increases with the size of the bank.

3.7 Payment Cards and Two-Sided Markets

Banks provide noncash payment services to their customers in the form of checks, transfers, direct debit systems, and payment cards. The payment card industry has grown substantially in the last decades. Invented in the 1950s in the United States (Evans and Schmalensee 1999), payment cards became incredibly successful, but only after they solved a ''chicken-and-egg'' problem. A payment card is only valuable to customers if it is accepted by sufficiently many retailers, and retailers find it profitable to accept a card only if sufficiently many consumers hold it. The challenge for the payment card platform (which can be an independent entity like American Express or an association of banks like Visa or MasterCard)[34] is to attract the two sides of the market while not losing money overall. Thus the payment card industry is a natural example of a "two-sided industry" (Rochet and Tirole 2003; 2006; Armstrong 2006), where competing platforms provide interdependent services to two (or several) categories of users. Examples of two-sided industries are software (Parker and Van Alstyne 2005), Internet portals (Caillaud and Jullien 2003), media (Anderson and Coate 2005), and intermediaries (Ellison, Fudenberg, and Möbius 2004; Jullien 2005). This section focuses on the payment card industry and shows how the traditional concepts of industrial organization and competition economics (monopoly outcome, Bertrand competition, social welfare) must be adapted to the analysis of indirect externalities in two-sided markets.

3.7.1 A Model of the Payment Card Industry

This model was developed by Rochet and Tirole (2003) and extended by Wright (2003), Guthrie and Wright (2003), and Rochet and Tirole (2006). It considers an economy where consumers can pay by cards or by an alternative means of payment (say, cash).

We present here a simplified version of the Rochet and Tirole model where the card network is for-profit and contracts directly with retailers. Cards are issued by competing banks. This fits well the case of a proprietary network such as American Express.[35] Figure 3.3 shows the costs and benefits attached to a card transaction.

All these costs and benefits are measured with respect to a cash payment. For example, b_B represents the cost saved by the buyer when he pays by card instead of cash. b_B is a random variable, realized at the time of purchase and only observed by the buyer. For example, if the buyer has no cash in his pocket (and not enough time to find an ATM), b_B may be high. The demand for card payments (measured by the fraction of purchases settled by card) is $D(p_B) = \Pr(b_B > p_B)$, where p_B is the buyer fee.

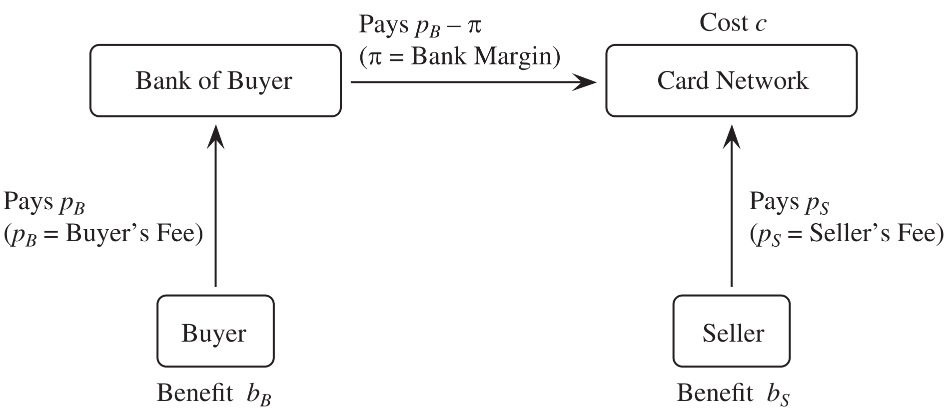

Figure 3.3
Costs and benefits of a card transaction.

Similarly, b_S represents the cost saved by the retailer when the payment is made by card instead of cash. We assume that b_S is the same for all transactions and observable by the card network. For simplicity, we neglect the cost incurred by the bank of the customer (the issuer of the card) when the payment is processed. The processing cost is entirely borne by the network. However, to compensate the bank of the buyer from incurring the cost of issuing the card, we assume the issuer collects a fixed margin π on each payment. Assuming that the total number of transactions (card + cash) is fixed and normalized to 1, we can compute the expected welfare of the different protagonists:

Consumer surplus $\quad u - p + \int_{p_B}^{\infty} D(s)\,ds,$

where u is the utility of consuming the good, p is the retail price, and the integral represents the option value associated with the possibility of paying by card:

$$\int_{p_B}^{\infty} D(s)\,ds = E[\max(0, b_B - p_B)].$$

Bank profit $\quad \pi D(p_B).$

Network profit $\quad (p_B + p_S - c - \pi)D(p_B).$

Seller profit $\quad p - \gamma + (b_S - p_S)D(p_B),$

where γ is the cost of the good for the seller.

3.7.2 Card Use

Card use depends on two things: whether retailers accept them, and how often customers want to use them. Retailers' decisions depend in turn on the price p_S they face (the merchant service charge) and on their competitive environment. Rochet and Tirole (2002) study retailers' card acceptance in the Hotelling-Lerner-Salop model, and Wright (2003) considers monopoly and Cournot models. In all cases, there is a maximum value of p_S above which merchants reject cards. We consider for simplicity the case of perfectly competitive retailers. In this case, retailers are segmented: some reject cards (and charge a lower retail price, equal to the marginal cost γ of the good); the rest accept cards but charge a higher[36] price (the increment being equal to the expected net cost $(p_S - b_S)D(p_B)$ of card payment). Since consumers are ex ante identical, they will choose the store accepting cards if and only if their option value for card payments exceeds this incremental price:

$$\int_{p_B}^{\infty} D(s)\,ds \geq (p_S - b_S)D(p_B).$$

This condition can be reformulated as

$$\phi \equiv \int_{p_B}^{\infty} D(s)\,ds + (b_S - p_S)D(p_B) \geq 0, \tag{3.35}$$

where ϕ represents the total surplus of final users, namely, buyers and sellers.

Condition (3.35) characterizes card use when only one network operates (monopoly). When several identical cards compete (Bertrand competition), only the ones such that ϕ is maximum are effectively used. This fundamental result holds true also when retailers have some market power (Rochet and Tirole 2002; Wright 2003).

3.7.3 Monopoly Network

When there is only one (for-profit) network, it selects the two prices p_B and p_S that maximize its profit:

$$B = (p_B + p_S - c - \pi)D(p_B) \tag{3.36}$$

under the constraint that cards are used

$$\phi \equiv \int_{p_B}^{\infty} D(s)\,ds + (b_S - p_S)D(p_B) \geq 0. \tag{3.37}$$

The Lagrangian of this problem is

$$L = B + \lambda\phi,$$

and the first-order conditions are

$$\frac{\partial L}{\partial p_B} = D(p_B) + (p_B + p_S - c - \pi)D'(p_B) - \lambda D(p_B) + \lambda(b_S - p_S)D'(p_B) = 0,$$

$$\frac{\partial L}{\partial p_S} = D(p_B) - \lambda D(p_B) = 0.$$

The second condition is equivalent to $\lambda = 1$. Then the first condition gives

$$p_B = c + \pi - b_S.$$

p_S is then determined by the constraint

$$\phi = 0 \Rightarrow p_S = b_S + \frac{\int_{p_B}^{\infty} D(s)\,ds}{D(p_B)}.$$

Thus $p_S > b_S$, which means that card payments increase the retailers' cost. This increase is passed on to consumers through an increase in retail prices. The social surplus generated by the card network is thus shared between the (monopoly) network and the banks.

3.7.4 Competing Payment Card Networks

If there are two (or more) networks that offer perfectly substitutable cards (Bertrand competition), only the ones that offer the maximum total user surplus ϕ will be used. The outcome of Bertrand competition is therefore characterized by prices p_B, p_S that solve

$$\begin{cases} \max \phi = \int_{p_B}^{\infty} D(s)\,ds + (b_S - p_S)D(p_B) \\ \text{under } p_B + p_S \geq c + \pi, \end{cases}$$

Namely, that maximize total user surplus under the break-even constraint of the platform. Since this constraint is clearly binding, we can write

$$p_S = c + \pi - p_B,$$

and thus

$$\phi = \int_{p_B}^{\infty} D(s)\,ds + (b_S - c - \pi + p_B)D(p_B).$$

This is maximum when

$$p_B = c + \pi - b_S,$$

and thus

$$p_S = b_S.$$

3.7.5 Welfare Analysis

By adding the welfares of all protagonists, we obtain the expression of social welfare:

$$W = [u - \gamma] + \int_{p_B}^{\infty} D(s)\,ds + (p_B + b_S - c)D(p_B).$$

Note that W does not depend on p_S. It is maximum for $p_B = p_B^W \equiv c - b_S$. Comparing with the formulas obtained for the monopoly and competitive cases, we see that

- both in the monopoly and competitive cases, buyer prices are higher than the welfare-maximizing level p_B^W;
- network competition leads to lower seller prices (and indirectly to lower retail prices) but does not change card use (and thus social welfare).

Thus in a two-sided industry, perfect competition between platforms does not necessarily lead to a social optimum. This is even more striking in the case where platforms are not-for-profit associations of banks (Rochet and Tirole 2002). In this case, the networks do not make any profit:

$$p_B + p_S = c + \pi,$$

and p_B is chosen to maximize the profit of banks $\pi D(p_B)$ under the constraint that $\phi \geq 0$ (so that cards are used effectively).

In the monopoly case, p_B is then the minimum buyer price such that

$$\phi = \int_{p_B}^{\infty} D(s) + (b_S - c - \pi - p_B) D(p_B) \geq 0,$$

which is lower than the optimal price p_B^W when

$$\int_{p_B^W}^{\infty} D(s)\, ds - \pi D(p_B^W) \geq 0,$$

that is, when the profit margin π of issuers is not too large.

The competitive case gives the same outcome as when networks are for-profit: Bertrand competition automatically reduces their profit to zero.

Therefore, not-for-profit associations tend to choose too low buyer prices when they have market power (monopoly) and too high buyer prices when they are in competition.

3.8 Problems

3.8.1 Extension of the Monti-Klein Model to the Case of Risky Loans

This problem is adapted from Dermine (1986). Modify the model of section 3.2 by allowing borrowers to default. More specifically, suppose that the bank has lent L to a firm that has invested it in a risky technology with a net (unit) return $\tilde{y}$. In the absence of collateral, the net (unit) return to the bank will be $\min(r_L, \tilde{y})$. When $\tilde{y} < r_L$, the firm defaults, and the bank seizes the firm's assets, which are worth $(1 + \tilde{y})L$.

1. Assuming that the bank has no equity (in conformity with the model of section 3.2), show that the bank itself will default if $\tilde{y}$ is below some threshold y^*. Compute y^*.

2. Assume risk neutrality and limited liability of the bank. The bank chooses the volumes L^* of loans and D^* of deposits that maximize the expectation of the positive