Interchange Fees and Inefficiencies in the Substitution between Payment Cards and Cash.

Marianne Verdier*

March 29, 2010

Abstract

This article explains why the collective determination of interchange fees in payment platforms and in ATM networks can lead to inefficiencies in the substitution between payment cards and cash. The merchant's bank, the acquirer, receives an interchange fee from the cardholder's bank, the issuer, each time a consumer pays by card, whereas the latter pays an interchange to the ATM owner when the consumer withdraws cash. If the issuers are ATM owners, I show that the divergence between the profit maximizing and the welfare maximizing interchange fees depends on the value of card payments and on the costs of cash.

JEL Codes: G21, L31, L42.

Keywords: Payment card systems, interchange fees, two-sided markets, money demand, ATMs.

*Université Paris Ouest Nanterre, Economix, Bâtiment K, 200 avenue de la République, 92001 Nanterre Cedex, France; E-mail: marianne.verdier@u-paris10.fr. I would like to thank Marc Bourreau, Joanna Stavins, Emilio Calvano, Jocelyn Donze and Isabelle Dubec for helpful discussions. I also thank conference participants for their remarks (IOC 2009, WEAI 2009, EARIE 2009, and AFSE congress 2009), and seminar participants at the Toulouse School of Economics.

1 Introduction:

Payment cards are widely hold and used in developed countries. In the European Union, for instance, in 2008, there were 1.46 payment cards per inhabitant. The increase in the number of payment cards has been accompanied by a rise in the number of Automatic Teller Machines (ATMs), which enable the consumers to use their payment cards to withdraw cash. Therefore, payment cards offer the consumers a convenient possibility to trade off between cash and card usage at the Point of Sales (POS), and empirical evidence shows that consumers often use their cards to withdraw cash (see Table 1). In 2007, in the Euro area, there were 880 million ATMs per inhabitant, and the total amount of cash withdrawals was estimated at 10% of the GDP.[1]

Table 1: Examples of card usage in European countries in 2006 (number in billion).[2]

	DE	FR	BE	SE	GB	IT
Total number of card transactions[3]	4.95	6.91	1	1.273	9.55	1.24
Percentage of cash withdrawals	49.9%	21.1%	25.7%	24%	28.8%	37.9%

An important economic issue is whether the level of card usage at the POS is socially optimal, and, if not, how to provide consumers with incentives to make efficient decisions when they substitute cards for cash. Several empirical studies (e.g Bergman et al. (2007)) prove that, in terms of social costs, cash is the most expensive payment instrument, whereas the use of debit cards is often too low to maximise social welfare.[4] Such inefficiencies arise because the consumers receive price signals that do not reflect the social costs of their payment choices. In particular, the consumers' private costs of using cash are rather low, as the use of cash is only charged when the consumers withdraw cash from the ATMs, and not at the POS.

The consumers' cost of using each payment instrument depends on complex cross-subsidization mechanisms. In Europe, banks often charge the use of payment instruments through the deposit fee when the consumers open an account. They also charge transaction fees that may be lower than cost - even sometimes negative- because of interbank transfers called "interchange fees". Banks use two different types of interchange fees: interchange fees on card payments and interchange fees on withdrawals.[5] On the one hand, interchange fees on card payments are

[1] Source: BIS statistics 2007.

[2] Source ECB Blue Book 2006, except for the United Kingdom APACS 2007.

[3] Number in billion of POS+withdrawal transactions proceeded in the country with a card issued in the country (all types of cards included).

[4] In Bergman et al. (2007) the per transaction cost of cash in Sweden is EURO 0.52, while the per transaction cost of debit cards is EURO 0.34.

[5] Agreements on Multilateral Interchange Fees are used in most payment systems. Notable exceptions include

paid by the merchant's bank (the acquirer) to the cardholders' bank (the issuer) each time a consumer pays by card. By lowering the cost of the issuer, interchange fees on card payments may contribute to lower the transaction fee that is paid by the cardholder, which might increase card usage. On the other hand, interchange fees on withdrawals are paid by the issuer of the card to the ATM owner each time a consumer withdraws cash. By raising the costs of the issuer, interchange fees on withdrawals may increase the transaction fee that the cardholder pays for withdrawals and reduce the number of cash withdrawals. Hence, at first sight, one could think that the existence of interchange fees encourages consumers to substitute cards for cash. However, banks may choose interchange fees that generate an inefficient level of substitution between cards and cash, as they do not internalize the consumer and the merchant surplus when they maximise their joint profit.

Given the recent decisions of Central Banks or Competition Authorities (e.g in Australia, in the European Union)[6] to regulate interchange fees, its seems particularly important to analyse whether a separate regulation of interchange fees in ATM networks and interchange fees in payment card systems can help to reach the socially optimal level of payment card usage.

This paper aims at studying if privately optimal interchange fees differ from the social optimum when the issuing banks are also ATM owners, and when the same organization manages the payment card system and the ATM network. For this purpose, I set up a model that takes into account how each agent that is involved in the payment process (bank, consumer, merchant) trades-off between the use of cash and debit cards. The main result of my paper is to show that the profit maximising interchange fee on card payments is too high to maximise the total user surplus (defined as the sum of the consumer and the merchant surplus) if the cost of cash is low for banks and for merchants, and if the value of the expenses paid by card is high. The profit maximising interchange fee on withdrawals is too high to maximise the total user surplus if the benefit of accepting cards is small for merchants, and if the cost of cash for banks is sufficiently low.

In my model, two issuing banks, which are also ATM owners, compete "à la Hotelling" on the market for deposits, after the choice of interchange fees for card payments and ATM transactions. Banks' ATMs are compatible, and each time a consumer withdraws cash from an ATM that is not owned by its bank, the issuer of the card pays to the ATM owner an interchange

the Netherlands, where discussions are ongoing with the NMa (competition authority) to authorize bilateral interchange fees.

[6] For the European Commission, see for instance the MasterCard decisions (IP/09/515 and IP/07/1959). As regards Australia, under the Payments Systems (Regulation) Act of 1998, the Reserve Bank of Australia has the power to regulate interchange fees and set standards, and it decided to exercise its power against MasterCard and Visa in 2006 and 2008.

fee. The ATM owner incurs fixed costs per transaction, and variable costs that depend on the volume of cash that is withdrawn from its network. On the merchant side, the acquirers are perfectly competitive and pay an interchange fee to the issuer of the card each time a consumer pays by card.

A consumer who opens an account in a bank is delivered a debit card, which enables him to pay at the POS and to withdraw cash from the ATMs. The consumer decides on where to establish an account by comparing the fees charged by each bank (deposit fee + transaction fees), as he anticipates that he will use the debit card during the next period to pay for his expenses and to withdraw cash. The consumer's transaction demand is modelled using the framework of Whitesell (1989)[7], who assumes that consumers make transactions of variable sizes, which can be paid either cash or by card. Given the costs of withdrawing cash, the costs of storing cash, the costs and the benefits of paying by card, a consumer chooses how much cash to withdraw from the ATM network, and the transaction value above which he pays by card, in order to minimize his transaction costs. The transaction costs born by a consumer depend the payment card fee and the withdrawal fee that are charged by his bank. The consumer can withdraw cash for free in an ATM that is owned by his bank and he has to pay a fee to his bank when he conducts a transaction outside of its bank's ATM network. He also has to pay a card fee when he uses his card.

On the merchant side, I consider homogenous merchants, who have to decide on whether or not to accept cards, depending on their cost of cash and on the merchant fee charged by the acquirers. I assume that the merchants are able to surcharge card payments and that there is imperfect pass through of the merchant fee to the consumers. With this framework, all merchants accept cards if the interchange fee on card payments is not too high.

I show that, in equilibrium, issuing banks price the transactions at their average perceived cost, as in Massoud and Bernhardt (2002) or Donze and Dubec (2006). The payment card transaction fee is equal to the issuer's marginal cost minus a subsidy that reflects the costs of cash for the bank. If the costs of cash are high, the payment card fee may even be negative. The withdrawal fee is equal to the average cost born by the issuer, which depends on the frequency of foreign withdrawals and on the ATM interchange fee. The intuition for this result is the following: Consumers internalize the expected transactions cost when they choose where to establish a bank account. When the transactions are priced at the average perceived costs, a consumer chooses to renounce to pay cash only if it increases the joint surplus that is obtained by its bank and by himself. Hence, banks can encourage efficient use of cash and cards when

[7]Empirical evidence of the Whitesell model has been provided by Raa and Shestalova (2004).

consumers make transactions at the following period, while extracting a part of their surplus through the deposit fee.

The deposit fee charged by a bank depends on the average surplus that a consumer obtains from opening an account and on the net opportunity cost for the bank of losing a "foreign" consumer when it attracts a new consumer. The novelty of my paper (compared to Massoud and Bernhardt (2002) or Donze and Dubec (2009) for instance) is that the the opportunity cost of losing a foreign consumer depends not only on the benefits that can be made on foreign consumers through foreign withdrawals, but also on the costs of cash for banks, as an issuer that also owns ATMs trades-off between:

- the benefit that it makes on "home" consumers through the deposit fee and the transaction fees,
- the costs of each payment instrument used by "home" consumers,
- the benefit that it makes on "foreign" consumers through the interchange fee when a "foreign" consumer makes a withdrawal in its network,
- the variable costs of cash generated by foreign withdrawals.

At the first period, the payment platform chooses the interchange fees that maximise banks' joint profit. As the acquirers are perfectly competitive, the payment platform takes only into account the issuers' profit. The profit-maximising interchange fees reflect a trade-off between the profits made on deposits and the profits made on foreign withdrawals, provided that merchants accept cards. As banks price the transactions at their average perceived cost, their joint profit depends on the deposit fee, which is determined by the opportunity cost of losing a foreign consumer, and on the costs of cash. When the interchange fees increase, the opportunity cost of losing a foreign consumer and the costs of cash are impacted as follows.

A rise in the interchange fee on card payments encourages the consumers to substitute cards for cash, if the surcharge rate is not too high. This reduces the opportunity cost of losing a foreign consumer, as consumers make fewer foreign withdrawals, and toughens the competition for deposits. This impact is negative for banks, but they may benefit from a reduction in the costs of cash, as cards are used more often as a substitute for cash. When the interchange fee on withdrawals increases, the profit per foreign withdrawal is increased, but the volume of foreign withdrawals falls, which reduces the costs of cash. Depending on how these two effects compensate each other, this may either increase or decrease the opportunity cost of losing a foreign consumer.

In general, a profit maximising payment platform chooses interchange fees that reflect this trade-off. If the costs of cash are equal to zero, the profit maximising interchange fee on withdrawals is equal to the monopoly price, whereas if the costs of cash are strictly positive, it is higher than the monopoly price. If the costs of cash are very high, the profit maximising interchange fee on card payments is chosen such that consumers pay for all their expenses by card, provided that cards are accepted by the merchants at this rate.

Then, I show that the profit maximising interchange fee on card payments is too high to maximise the total user surplus if the cost of cash is low for banks and for merchants, and if the value of card payments is high. If the cost of cash is high for banks, the profit maximising interchange fee on card payments may be too low to maximise the total user surplus. I also show that the profit maximising interchange fee on withdrawals is too high to maximise the total user surplus if the benefit of accepting cards is small for merchants, and if the cost of cash for banks is not too high. The main policy implication of this result is that payment card services should not be analysed separately from cash provision services, because there are substitution effects between cash and cards, especially if the payment platform organizes the interactions between the issuers, the acquirers and the ATM owners.

My paper is related to two different strands of the literature on payment cards: the literature on interchange fees in payment systems and the literature on ATMs.[8] However, the relationship between optimal interchange fees in payment card systems and interchange fees in ATM networks has never been analysed in previous research papers.

The literature on interchange fees in payment card systems studies the divergence between the profit maximising and the welfare maximising interchange fees.[9] My paper departs from this literature by using Whitesell(1989)'s assumptions to model the consumer's demand for transactions. This way of modelling the consumer's demand enables me to relate the switching point between cash and card payments to the levels of interchange fees in ATM networks and in payment card platforms. This choice has been motivated by the empirical observation that, in many countries,[10] the same card can be used to pay at the POS and to withdraw cash, and

[8] In this paper, we do not study the determination of credit card fees. See Chakravorti and Emmons (2003), Chakravorti and To (2007) or Chakravorti and Bolt (2008). For a general survey on the economics of credit cards, debit cards and ATMs see Scholnick et al. (2008).

[9] For a review of the literature, see Rochet (2003) or Verdier (2010). According to Baxter (1983), interchange fees in payment card systems solve the usage externalities that arise when the consumers make the optimal choice of a payment instrument at the POS. The optimal interchange fees for card payments depend in particular on the nature of the strategic interactions between merchants (Rochet and Tirole (2002), Wright (2004)), on the nature of competition between banks (Rochet and Tirole (2002)), on the ability of the payment platform to surcharge (Wright (2002)), on banks' investments in quality (Verdier (2010)), and on the existence of competing payment platform (Rochet and Tirole (2003), Guthrie and Wright (2007), Chakravorti and Roson (2006)).

[10] See for instance Table 1 of Appendix J for examples of ATM and payment platforms in various european countries.

by the observation that issuers are often ATM owners.[11]

The other branch of the theoretical literature studies the welfare effects of interchange fees in ATM networks.[12] As shown by Matutes and Padilla (1994), the role of interchange fees in ATM networks is to provide competing banks with incentives to share their ATMs. My model builds on this result by assuming that banks' ATMs are already compatible, which justifies the use of an interchange fee on withdrawals. The literature also highlights two potential negative welfare effects due to the presence of interchange fees in ATM networks. First, Massoud and Bernhardt (2002) or Donze and Dubec (2006) show that interchange fees soften the competition on deposits, because it becomes less profitable to attract a consumer when a "foreign" consumer makes withdrawals that generate revenues. This first effect is also present in my model, and the literature shows that it is reinforced by the presence of "foreign fees", which I do not take into account in this article (see Massoud and Bernhardt (2002), or Donze and Dubec (2009)). A second negative welfare effect of interchange fees is related to an excessive deployment of ATMs, as shown by Donze and Dubec (2006). I discuss this issue in the extension section.

The rest of the article is organized as follows. In section 2, I start by presenting the model and the assumptions. In section 3, I solve for the equilibrium of the game and determine the levels of interchange fees that maximise banks' joint profits and that maximise social welfare. In section 4, I study the robustness of the results obtained in section 3 by examining the impact of merchants' heterogeneity, asymmetric issuers, issuers as acquirers, and by discussing the problem of endogenous ATM deployment. Finally, I conclude.

2 The model

The payment system is modelled as an association of issuing and acquiring banks, in which the issuers are also ATM owners.[13] The banks provide payment card services which allow their consumers to pay by card at the Point of Sales (POS) and to withdraw cash from the ATM network. Each time a consumer pays by card, an interchange fee is paid by the Acquirer of the transaction to the Issuer of the card, whereas each time a consumer withdraws cash, an interchange fee is paid by the Issuer of the card to the ATM owner. The payment system chooses the level of interchange fees for card payments and cash withdrawals that maximise

[11] Also, there is empirical evidence that consumers trade-off between cards and cash according characteristics of the transaction (see Bounie and François (2006), Borzegowski, Kiser and Ahmed (2008), Schuh and Stavins (2009)). Among these characteristics, the consumers take into account the value of the transaction when they choose their payment method: consumers tend to pay cash transactions of small value and pay by card transactions of larger amounts.

[12] For a review of this literature, see MacAndrews (2003).

[13] The model will be extended to the case in which the Issuers are also Acquirers in Section 4.

banks' joint profit.

Banks: Two issuing banks, denoted by 1 and 2, are located at the extremities of a linear city of length one and compete on the market for deposits. Each bank proposes a package of account services at a price P_i, where $i \in \{1; 2\}$. This package comprises the provision of a debit card,[14] which enables the consumers to pay by card at the POS, and to withdraw cash from the ATMs. As banks' ATMs are compatible, the consumers are allowed to withdraw cash from the ATMs that are not managed by the bank in which they hold an account (their "home" bank). A bank charges its consumers a price, w_i, when they withdraw cash from a foreign ATM, whereas "home" withdrawals are free. In the main model of this paper, I assume that "foreign fees" on withdrawals are not allowed, such that a bank does not charge foreign consumers when they use its ATMs.[15] As regards card transactions, banks charge their customers a price each time they pay by card, which I denote by f_i for all $i \in \{1; 2\}$.[16]

Banks incur fixed and variable costs when consumers withdraw cash and pay by card. I assume that a withdrawal transaction costs c_W to the ATM owner. Also, banks have to bear variable costs, which depend on the volume of cash that is withdrawn from their ATMs, such that if the volume of transactions in bank i's ATMs is V_i, bank i incurs the cost kV_i, where $k \in (0; 1)$. Each time a consumer uses its card for a payment, the Issuer of the card bears the cost c_I of providing the payment card service.[17]

In the main model, the merchants' banks, the Acquirers, are assumed to be perfectly competitive. The marginal cost of acquiring a card transaction is denoted by c_A.

Consumers Consumers are uniformely located along the linear city and may open an account either at bank 1 or at bank 2. They incur a linear transportation cost $t > 0$ per unit of distance when they travel to open an account. If a consumer decides to open an account, he obtains a benefit $B > 0$, which I assume to be sufficiently large, such that the market is covered. I also

[14] In this paper, the fact that the payment card may be used to obtain credit is not taken into account.

[15] In the literature on ATMs, a surcharge is the price that is charged by a bank to its consumers, when they withdraw cash from the ATMs of the other bank. A foreign fee is a price that is charged by a bank to foreign customers when they use its ATMs. In the literature on payment cards, merchants are said to surcharge if they are allowed to choose a higher price if the consumer pays by card. In this paper, I assume that there are no "foreign fees" charged by a bank when a foreign consumer withdraws cash - which is generally the case in Europe- but that merchants are able to surcharge card payments if $\beta > 0$ and that there are surcharges on withdrawals.

[16] Shy and Wang (2010) have shown that payment platforms earn more profit by charging proportional fees on both sides of the market rather than fixed fees. In this paper, I assume that issuing banks charge fixed fees, while acquiring banks charge proportional fees.

[17] In this analysis, I neglect the fixed costs of setting up the payment card infrastructure. According to De Grauwe et al. (2006), who compare studies conducted by the Central Banks in the Netherlands and in Belgium, the fixed costs of cash amount to 40% of the total cost of cash in both countries. The variable costs of cash that depend on the number of transactions amount to 40% of the costs of cash, while the variable costs of cash that depend on the transaction value make about 20%.

assume that it is never in the interest of a consumer to open an account in both banks. Once he has opened an account, the consumer owns a debit card, which enables him to pay at the POS, or to withdraw cash from ATMs. For all $i \in \{1;2\}$, I assume that a consumer of bank i makes an exogenous percentage $\varphi_i \in [0,1]$ of withdrawals in the ATMs of his bank ("home" ATMs) and a percentage $1-\varphi_i$ of withdrawals in "foreign" ATMs .[18]

To model the consumers' demand for transactions, I use the framework of Whitesell (1989), in which the consumers make transactions of variable sizes. The size of a transaction is denoted by T, where T belongs to $\left[0,\overline{\lambda}\right]$. Transactions of each size are assumed to occur at a uniform rate over a unit period. If $F(T)$ represents the value of spending on all transactions of size T during the period, total spending is given by[19]

$$S = \int_0^{\overline{\lambda}} F(T)dT.$$

A consumer obtains a surplus V from his purchases, which is assumed to be sufficiently large, such that the consumers always bear the transaction costs needed to spend S.

I make the following assumptions on F to ensure that there exists an equilibrium in which consumers use both cash and cards to pay for their expenses:

(A1) F is twice differentiable over $\left[0,\overline{\lambda}\right]$.

(A2) F is increasing and concave over $\left[0,\overline{\lambda}\right]$.

(A3) $\lim\limits_{\lambda \to 0} \dfrac{\lambda}{\sqrt{\int_0^{\lambda} F(T)dT}} = l$, with l belonging to $\mathbb{R}$.

As in Whitesell (1989)'s model, the problem of a consumer is to decide which transactions to pay cash, and the amount of cash to hold and to withdraw from the ATM network. The consumers incur fixed and variable costs and benefits, which differ if they use cash or if they pay by card. If a consumer pays by card, he has to pay the fixed fee f_i to its bank, and but he obtains a variable net benefit $v_i > 0$, which depends on the size of the transaction.[20] The

[18] The APACS report "The way we pay" (2008) shows that, for instance in the United-Kingdom, asymmetries between banks are common place as regards the percentage of "on-us" transactions. In this model, banks have a part of captive "foreign" customers, who need to withdraw cash, and only find an ATM that is managed by the other bank. This assumption is relaxed in the extension Section, in which I give intuitions of the results obtained when the probability to withdraw from "home" ATMs depends on banks' deployment strategies.

[19] The transaction costs are not included in the total volume of spending.

[20] The variable net benefit paying by card v_i depends on the bank where the consumer holds an account. We allow for some differentiation between the banks for the "card payment" service, but for simplicity, this differentiation is assumed to be exogenous.

net benefit v_i can be interpreted as the insurance services or the rewards, which depend on the size of the transaction, net of the transaction costs.[21] As there is an opportunity cost $r > 0$, associated to the detention of cash[22], the consumers may decide to make several withdrawals to obtain the amount of cash needed to pay for their expenses. The number of withdrawals made by the consumers of bank i is denoted by n_i. I also assume that the consumers bear an exogenous fixed cost $b > 0$ when they withdraw cash, which can be interpreted as the time needed to find an ATM.

Given this cost structure, a consumer who has an account at bank i decides to pay by card if the value of the transaction exceeds some threshold λ_i, where λ_i belongs to $\left[0, \overline{\lambda}\right]$. The consumers of each bank i decide on the optimal value of the threshold λ_i and on the number of withdrawals n_i so as to minimize their transaction costs, which are denoted by C_i, for all $i \in \{1; 2\}$.[23]

Merchants All merchants have to bear variable costs when the consumers use payment instruments. They have the choice between accepting and refusing payment cards. If they accept payment cards, each time a consumer pays by card, they pay a fee to their bank, $M(T)$, which is assumed to depend of the size of the transaction, such that $M(T) = mT$, with $m > 0$.[24] The merchants are also assumed to pass through the cost of the merchant fee to the consumers who pay by card at a rate β, which is assumed to be small. Hence, if the consumers pay by card a transaction of value T, they have to pay a surcharge of βmT to the merchant. When the consumers pay cash, the merchants have to bear the variable costs of collecting cash, counting the notes and the coins, and carrying them to their banks. The variable costs of cash payments for all merchants is denoted by c_M.[25]

[21] Variable costs and benefits could also depend on other characteristics of the transaction, such as the spending place or the type of good which is purchased. Bounie and François (2006) investigated empirically the determinants of the use of payment instruments at POS. They found strong evidence of the effect of the transaction size on the choice of the payment instrument. The other variables that influences significantly the choice of the payment instrument are: the type of good and the spending place, the restrictions on the supply-side and the organization of the payment process. Boeschoten (1998) also demonstrates the importance of the transaction size.

[22] In this model, I assume that r is not a strategic variable that can be decided by the bank which manages the deposit account. This opportunity cost is similar to Baumol (1952)'s model of money demand.

[23] My model differs from the main model of the literature on payment cards (See Verdier(2010) for a survey of this literature) as consumers' heterogeneity plays a role for the competition on the market for deposits.

[24] This assumption is consistent with the industry practices. According to Arango and Taylor (2008) who surveyed the merchants' costs of accepting payment instruments in Canada, for a transaction of \$36.5, the cost of cash is \$0.25 and the cost of debit cards is \$0.19. Cash becomes less costly for transactions under \$12.

[25] As the model is already complex, I decided not to assume that the merchants are heterogeneous over c_M or over β as in Chakravorti and Bolt (2008). The assumption that merchants are homogenous is relaxed in the extension Section. In my analysis, this cost is considered as exogenous. However, if the merchants' banks are perfectly competitive, this cost could reflect partially the price that the merchants have to pay to their banks when they deposit cash at their bank's branch. Also, in this paper, I do not consider the effects of the strategic interactions between merchants. Rochet and Tirole (2009) argue that merchants may be ready to accept cards

Payment system: The payment system organizes the interactions between the banks by choosing the level of the interchange fee on card payments, a^C, and the interchange fee on cash withdrawals, a^W, so as to maximise banks' joint profit. The interchange fee on card payments is paid to the Issuer of the card by the Acquirer of the transaction each time the consumer pays by card. The interchange fee on cash withdrawals is paid by the Issuer of the card to the ATM owner, and is assumed to be higher or equal to the marginal cost c_W.[26] The interchange fees are assumed to be positive and to be paid on a "per-transaction" basis. I also assume that if merchants refuse cards, banks have to bear a desutility $Z > 0$, which is sufficiently large, such that it is never in the interest of the banks to choose interchange fees which encourage merchants to refuse cards.[27]

I also make the following assumptions:

(A4) For all $i \in \{1, 2\}$, $c_I \leq \overline{\lambda}(v_i + \sqrt{r(c_W + b)/2S(\overline{\lambda})})$.

(A5) $\dfrac{c_M}{1-\beta} \leq c_I + c_A$.

(A6) For all $i \in \{1, 2\}$, $\beta \leq \dfrac{v_i}{v_i + c_M}$.

(A7) The level of interchange fees has no impact on the distribution of the goods' prices, F. However, it impacts the transaction costs born by the consumers and the merchants.

Assumption (A4) is verified if the variable card benefit, v_i, is high enough. It ensures that the consumers do not use only cash to pay for their expenses if a withdrawal transaction is priced at the marginal cost of the ATM owner, c_W, and if a card payment is priced at the marginal cost of the issuer, c_I.[28]

Assumption (A5) is standard in the literature. As we will see in our analysis, it is necessary to ensure that the consumers do not pay by card all their purchases.

Assumption (A6) ensures that the surcharge rate is not too high, such that, in equilibrium, the variable benefit of paying by card for consumers is always higher than the surcharge than they must paid to the merchant.

---

even if the fee they have to pay exceeds their convenience benefit, as they internalize a fraction of the consumer's surplus in their decision to accept cards. This effect leads to an increase in the maximum interchange fee that is compatible with merchants' acceptance of payment cards.

[26] Otherwise, it would not be profitable for banks to invest in ATM deployment or to reach full compatibility (See Matutes and Padilla (1994)).

[27] Several explanations can justify the existence of a desutility for the banks when the merchants refuse payment cards. Firstly, ATM owners have to bear large variable costs of handling cash, which can become excessive if consumers pay cash for all their expenses. Secondly, the value of holding the payment card decreases for the consumers. Hence, banks can lose consumers if payment cards are refused by the merchants. In Appendix D-5, I will provide ex post the conditions on Z such that it is never in their interest that merchants refuse cards.

[28] I will show in the proof of Proposition 1 that the right side of the inequality represents the average cost of cash, if the consumer pays all his expenses cash and if the withdrawals are priced at the marginal cost of the ATM owner.

Assumption (A7) means that the level of interchange fees does not impact the retail prices. Therefore, interchange fees impact the consumers' choices only through the prices of the payment instruments. Empirical studies have shown that the links between the level of interchange fees and retail prices are difficult to measure. In Australia, for instance, a fall in the level of interchange fees has not triggered a reduction of retail prices.[29] Lifting this assumption would not change the intuitions of the results obtained in my paper.

Timing: The timing of the game is as follows:

1. The payment platform chooses the interchange fee for card payments, a^C, and the interchange fee for cash withdrawals, a^W.
2. The issuing banks choose the fees P_i, f_i, and w_i and the acquirers choose the merchant fee m.
3. Merchants decide whether or not to accept payment cards.
4. Consumers choose the bank from which to hold an account and a payment card.
5. Consumers make their purchases. They choose the number of cash withdrawals, and the threshold which separates card and cash payments.

In the following section, I look for the subgame perfect equilibrium, and solve the game by backward induction.

3 The equilibrium:

3.1 Stage 5: payments and withdrawals decisions.

In this section, I study the consumers' payment and withdrawal decisions. At the last stage of the game, the consumer already holds a debit card, which is issued the bank in which he has an account. He has to choose how to pay for a total amount of expenses, S, in order to minimize its transaction costs, as in Whitesell (1989).

3.1.1 The consumer's payment decisions if merchants accept cards.

The number of withdrawals and the threshold value for card payments. I start by assuming that the merchants accept cards and I study the consumers' payment and withdrawal

[29] See Chang, Evans, and Swartz (2005).

decisions. The costs of the payment instruments consist of the fixed costs paid for each transaction, the variable benefit of paying by card, the opportunity cost of cash detention, and the costs of cash withdrawals. A consumer pays cash if the transaction amount T belongs to $[0, \lambda_i]$, and pays by card if T belongs to $\left[\lambda_i, \overline{\lambda}\right]$. From the assumptions on F, the value of the transactions that are paid cash, $S(\lambda_i)$, is given by

$$S(\lambda_i) = \int_0^{\lambda_i} F(T)dT,$$

whereas the value of the transactions that are paid by card is given by

$$S - S(\lambda_i) = \int_{\lambda_i}^{\overline{\lambda}} F(T)dT.$$

I now precise the total costs born by a consumer for his payment transactions. There are $F(T)/T$ transactions of size T. A consumer pays cash if T belongs to $[0, \lambda_i]$ and pays by card otherwise. As he obtains a variable benefit v_i of paying by card, while paying a surcharge to the merchant, the net costs of card payments are

$$f_i \int_{\lambda_i}^{\overline{\lambda}} \frac{F(T)}{T} dT + (\beta m - v_i)(S - S(\lambda_i)),$$

where f_i is the transaction fee that the consumer pays to his bank.

Finally, the consumer has to bear the costs of withdrawing and holding cash. In average, if $n_i > 0$, the consumer holds a quantity $S(\lambda_i)/(2n_i)$ of cash in his pocket, so the opportunity cost of cash detention is $rS(\lambda_i)/(2n_i)$, as in Baumol (1952) or Tobin (1956). Each time the consumer goes to an ATM, he bears a fixed exogenous cost b. "Home" withdrawals are free, but the consumer pays the price w_i to his bank for "foreign" withdrawals, which happens in $(1 - \varphi_i)$% of the cases, so the total cost of cash withdrawals is $n_i^*((1 - \varphi_i)w_i + b)$. To sum up, if $n_i > 0$, the costs of withdrawing and holding cash are

$$\frac{r}{2n_i} S(\lambda_i) + n_i((1 - \varphi_i)w_i + b).$$

If $n_i > 0$ and λ_i belongs to $\left[0, \overline{\lambda}\right]$, I can express the total transaction costs of a consumer that holds an account at bank i as a function of λ_i and n_i, that is

$$C_i(\lambda_i, n_i) = \frac{r}{2n_i} S(\lambda_i) + n_i((1-\varphi_i)w_i + b) + f_i \int_{\lambda_i}^{\overline{\lambda}} \frac{F(T)}{T} dT + (\beta m - v_i)(S - S(\lambda_i)). \quad (1)$$

The consumer determines the optimal number of cash withdrawals, n_i^*, and the optimal value of the transaction, λ_i^*, which minimize its total transaction costs, that is, $C_i(\lambda_i, n_i)$. The following proposition summarises the results, which are similar to Whitesell (1989).

Proposition 1 *Assume that cards are accepted by merchants. If* $f_i > l\sqrt{r((1-\varphi_i)w_i + b)/2}$ *and if* f_i *is not too high compared to the average cost of using only cash, there exists a unique transaction value* $\lambda_i^* \in (0, \overline{\lambda})$ *above which the consumer of bank* i *pays his expenses by card. If* $f_i \leq l\sqrt{r((1-\varphi_i)w_i + b)/2}$*, the consumer pays all his expenses by card. If the card fee is sufficiently high compared to the average cost of using only cash, the consumer does not pay by card.*

Proof. See Appendix A-1. ■

Consumers trade off between cash and the payment card at the POS. Proposition 1 shows that, if the card fee is not too low, a consumer pays by card if the amount of the transaction is high, and pays cash otherwise. This is because the variable benefit of paying by card, v_i, is higher for transactions of larger amounts. Also, the cost of holding cash ($rS(\lambda_i)/2n_i$) increases with the value of the expenses that are paid cash.

From Appendix A-1, if $f_i > l\sqrt{r((1-\varphi_i)w_i + b)/2}$, the optimal number of withdrawals, n_i^*, is given by

$$n_i^* = \sqrt{\frac{rS(\lambda_i^*)}{2((1-\varphi_i)w_i + b)}}.$$

The optimal number of withdrawals is expressed as in Baumol (1952)'s model, except that the volume of transactions that is paid cash, $S(\lambda_i^*)$, depends on the trade-off that consumers make between cash and the payment card, as in Whitesell (1989).[30] Notice that, in my model, consumers pay all their expenses by card if the card fee is sufficiently low, as I consider that the consumer behavior depends only on the costs of the payment instruments.[31]

Some comparative statics.

[30] See Appendix A-2 for an example with the function $F(T) = 2ST$, with T belonging to $[0, 1]$.

[31] In this analysis, we neglect the other attributes of the payment instruments that may be valued by the consumers. For instance, Chakravorti and Bolt (2008) build a model in which consumers participate in payment card networks to insure themselves against three types of shocks: income, theft, and the type of merchant they are matched to. Another motive for using cash would be the fact that cash payments are anonymous.

Lemma 1 *The optimal threshold, λ_i^*, and the number of withdrawals, n_i^*, increase with the card fee, and increase with the merchant fee if $\beta > 0$. They decrease with the withdrawal fee and with the variable benefit of paying by card.*

Proof. See Appendix A-3. ■

When the card fee decreases, a consumer chooses more often to pay by card, and withdraws cash less frequently. If the merchant fee decreases, the surcharge paid by the consumer is reduced, and the consumer also pays more often by card. Similarly, when the variable benefit of paying by card becomes higher, the transaction value above which consumers pay by card is reduced, whereas the number of cash withdrawals decreases.

Now that I have determined the optimal usage of payment instruments, I can express the total cost that is born by a consumer as a function of λ_i^*. If the consumer uses both payment instruments, that is, if λ_i^* belongs to $\left(0, \overline{\lambda}\right)$, we have at the optimum,

$$C_i^*(n_i^*, \lambda_i^*) = C_i^*(\lambda_i^*) = \sqrt{2r((1-\varphi_i)w_i + b)S(\lambda_i^*)} + f_i \int_{\lambda_i^*}^{\overline{\lambda}} \frac{F(T)}{T} dT + (\beta m - v_i)(S - S(\lambda_i^*)).$$

If the consumer pays all his transactions by card, we have that $\lambda_i^* = 0$, and $n_i^* = 0$, and the consumer's costs are

$$C_i^*(0,0) = C_i^*(0) = f_i \int_0^{\overline{\lambda}} \frac{F(T)}{T} dT + (\beta m - v_i)S.$$

Finally, if the consumer pays cash all his expenses, we have that $\lambda_i^* = \overline{\lambda}$, and $n_i^* = \sqrt{rS/2((1-\varphi_i)w_i + b)}$, and the consumer's costs are

$$C_i^*(n_i^*, \overline{\lambda}) = C_i^*(\overline{\lambda}) = \sqrt{2r((1-\varphi_i)w_i + b)S}.$$

In Lemma 2, I explain how the transaction costs of the consumer vary with the transaction prices charged by the banks and with the variable benefits of paying by card.

Lemma 2 *The consumer's payment costs, C_i^*, increase with the withdrawal fee, the card fee and the merchant fee, but decrease with the variable benefit that a consumer obtains from paying by card.*

Proof. See Appendix A-4. ■

Comparison with the switching point that minimizes the costs born by the users. In Lemma 3, I compare the threshold that separates card and cash payments with the threshold that would minimize the costs born by the consumer and the merchant.

Lemma 3 *If $m < c_M/(1-\beta)$, the consumer's switching point between card and cash payment exceeds the level that minimizes the total user cost.*

Proof. See Appendix A-5. ■

As the consumer does not internalize the merchant's cost of accepting each payment instrument, he may choose a switching point between card and cash payments that is not optimal from the point of view of the "joint user" (consumer+merchant). If $m < c_M/(1-\beta)$, the cost of accepting cards for merchants is lower than the cost of accepting cash. Hence, the switching point between card and cash payments exceeds the level that minimizes the total user cost. The fact that the consumer chooses its payment method without internalizing the merchant's acceptance costs is the first source of inefficiency in the substitution between cash and card payments.[32]

3.1.2 The consumer's payment decisions if the merchants refuse cards.

Finally, I study the consumer's payment decisions if the merchants refuse cards. If the merchants refuse payment cards, a consumer pays cash for all his expenses. The optimal number of withdrawals is similar to Baumol's model, that is $n_i^* = \sqrt{rS/2((1-\varphi_i)w_i + b)}$, and the consumers's costs are $C_i^* = \sqrt{2r((1-\varphi_i)w_i + b)S}$. Now that I have expressed the transaction costs that are born by the cardholders, I study their decision to open an account at stage 4.

3.2 Stage 4: Choice of the bank.

At stage 4, prior to making transactions, consumers have to decide on opening an account either at bank 1 or at bank 2. When they make their affiliation decision, consumers take into account the expected transaction costs at stage 4, the fixed deposit fee P_i, and the transportation cost, which depends on their location. A consumer located at point $x \in [0;1]$, that opens an account at bank i located at d_i, bears a cost $S + t\,|x - d_i| + P_i + C_i^*(\lambda_i^*)$, and obtains a surplus $V + B$. The marginal consumer is given by:

$$x = \frac{1}{2} + \frac{1}{2t}(P_2 - P_1 + C_2^*(\lambda_2^*) - C_1^*(\lambda_1^*)). \tag{2}$$

[32] As we will see in section 3.3, homogenous non strategic merchants always accept cards when it is less costly for them to do so. But if merchants were strategic, they would accept cards for higher values of the merchant fee than the cost of accepting cash. In this case, the consumers would be able to choose a switching point between cards and cash that is below the level that minimizes the total cost of users.

The market share of bank 1 is equal to $\gamma_1 = x$, whereas the market share of 2 is given by $\gamma_2 = 1 - \gamma_1$, provided no firm corners the market. Banks compete on the market for deposits on the total level of costs that they offer to their consumers, which depends on the price of deposits and on the transaction prices. The bank that offers the lowest level of costs to the consumers has the highest market share.

3.3 Stage 3: Card acceptance decision.

At stage 3, the merchants decide whether or not to accept cards. The merchants accept cards if their profit is higher when some consumers pay by card. As there are no strategic interactions between merchants, the card acceptance decision is determined by studying the behavior of a representative merchant. If the consumers pay cash for all their expenses, the representative merchant's profit depends on the value of the expenses minus the cost of cash,[33] that is

$$\Pi^{cash} = S - c_M S.$$

If some consumers pay by card, the representative merchant's profit depends also on the costs of card payments, that is

$$\Pi^{card} = S - c_M S(\lambda_i^*) - m(1-\beta)(S - S(\lambda_i^*)).$$

The merchants accept cards if $\Pi^{card} \geq \Pi^{cash}$, that is if $(c_M - m(1-\beta))(S - S(\lambda_i^*)) \geq 0$. As $S - S(\lambda_i^*) \geq 0$, the card acceptance condition is

$$c_M - m(1-\beta) \geq 0.$$

Hence, merchants accept cards if the merchant fee is not too high.

3.4 Stage 2: Bank fees.

In this section, I determine how banks price their account services, provided that the merchants accept cards. I start by determining the conditions under which merchants accept cards. Then I determine the deposit price and the transaction prices chosen by the issuers. Finally, I analyse how the prices affect the consumers' payment decisions at stage 5.

[33] The case in which merchants are heterogeneous will be discussed in the extension section.

3.4.1 The merchant fee.

First, I determine the conditions under which merchants accept payment cards, as this will influence the price that the issuers are able to choose when they compete for deposits. If the merchants accept cards, as the acquirers are perfectly competive, the merchant fee is equal to the acquirers' perceived marginal cost, which is the sum of the interchange fee paid to the Issuer of the card, and the acquisition cost, that is $m = a^C + c_A$. As I assumed that it is never in the interest of the banks that the merchants refuse cards,[34] from now on, I restrict the analysis to the subgame in which merchants accept payment cards. The merchants accept payment cards if

$$c_M - (1-\beta)(a^C + c_A) \geq 0,$$

which means that the interchange fee on card payments is bounded from above, that is $a^C \leq \overline{a^C}$, where $\overline{a^C} = c_M/(1-\beta) - c_A$.[35]

3.4.2 The transaction fees and the deposit fee.

At stage 2, each issuing bank chooses the deposit price P_i, and the transaction prices, f_i and w_i, that maximise its profit,

$$\pi_i = \gamma_i(P_i + M^i_{HC}) + (1-\gamma_i)M^i_{FC}, \tag{3}$$

where M^i_{HC} denotes the margin made on the transactions of a consumer that holds an account at bank i (a "home" consumer), whereas M^i_{FC} denotes the margin made on the transactions of a consumer that holds an account at bank j through foreign withdrawals (a "foreign" consumer).[36] The consumers of each bank can either use a combination of cash and card payments, or pay for all their expenses by card, or use only cash.

Let me detail here the components of the margin that bank i makes on "home" consumers' transactions. The margin made on "home" consumers' transactions comprises the price of payment card transactions, the interchange fee that is collected from the Acquirer for each transaction, and the price of "foreign" withdrawals, which is perceived for $(1-\varphi_i)\%$ of the withdrawals that are made by the consumers. The margin made on "home" consumers' transactions also

[34] The conditions under which this Assumption is verified at the equilibrium is given in Appendix D-5.

[35] Notice that, because of Assumption (A6), we can check that, for all $a^C \leq \overline{a^C}$, we have $v_i - \beta m > 0$, which was a condition to establish Proposition 1 and Lemma 1. This condition is equivalent to $\beta < v_i/(a^C + c_A)$ for all $a^C \leq \overline{a^C}$. It is sufficient to check that the condition is valid for $a^C = \overline{a^C}$, which is true as $\overline{a^C} + c_A = c_M/(1-\beta)$, by Assumption (A6).

[36] I provide in Appendix B-2 the conditions under which the second-order conditions of profit maximisation are verified.

involves the marginal costs of card payments, and the marginal costs of withdrawals, which differ if the consumer makes a "foreign" withdrawal, as the bank has to pay an interchange fee. Finally, the bank has to bear the variable cost of the volume of cash that is withdrawn by "home" consumers from the bank's ATMs, that is, $k\varphi_i S(\lambda_i^*)$.[37] Hence, we have:

$$M_{HC}^i = (f_i + a^C - c_I)\int_{\lambda_i^*}^{\overline{\lambda}} \frac{F(T)}{T}dT + n_i^*((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W) - k\varphi_i S(\lambda_i^*). \quad (4)$$

I now detail the components of the margin that bank i makes on "foreign" consumers' transactions.[38] The margin on "foreign" consumers comprises the profit obtained on "foreign" withdrawals through the interchange fee, and the cost of cash that is withdrawn from bank i's ATMs by "foreign" consumers, that is $k(1-\varphi_j)S(\lambda_j^*)$. Hence, we have:

$$M_{FC}^i = n_j^*(1-\varphi_j)(a^W - c_W) - k(1-\varphi_j)S(\lambda_j^*). \quad (5)$$

Proposition 2 gives the equilibrium deposit fee P_i^*, and the equilibrium transaction fees, f_i^* and w_i^*, that are chosen by each bank at stage 2. In this Proposition, I denote by $\left(M_{HC}^i\right)^*$ and $\left(M_{FC}^i\right)^*$ the margin that bank i makes at the equilibrium of stage 2 on the transactions made by "home" and "foreign" consumers, respectively.

Proposition 2 *Assume that cards are accepted by merchants. At the equilibrium of stage 2, banks price the transactions at the average perceived cost, that is* $f_i^* = c_I - a^C - \varphi_i k\lambda_i^*$ *and* $w_i^* = a^W + \varphi_i c_W/(1-\varphi_i)$. *The deposit fee is*

$$P_i^* = t + \left[2\left(M_{FC}^i\right)^* + \left(M_{FC}^j\right)^* - 2\left(M_{HC}^i\right)^* - \left(M_{HC}^j\right)^* + C_j^*(\lambda_j^*) - C_i^*(\lambda_i^*)\right]/3.$$

Proof. See Appendix B-1 and B-2. ■

Corollary 1 *At the equilibrium of stage 2, if cards are accepted by merchants, for each bank* $i \in \{1;2\}$, *the margin on home consumers is*

$$\left(M_{HC}^i\right)^* = -k\varphi_i\lambda_i^*\int_{\lambda_i^*}^{\overline{\lambda}} \frac{F(T)}{T}dT - k\varphi_i S(\lambda_i^*),$$

[37] By Assumption, if the volume of cash withdrawals from the ATMs of bank i is V_i, bank i bears a cost kV_i.

[38] Notice that our analysis is still valid if the consumers use either only cash or only cards to pay for their expenses.

whereas the margin on foreign consumers is

$$\left(M_{FC}^{i}\right)^{*}=n_{j}^{*}(1-\varphi_{j})(a^{W}-c_{W})-k(1-\varphi_{j})S(\lambda_{j}^{*}).$$

Each issuing bank that also owns ATMs trades off between the revenues obtained from "home" consumers, and the revenues obtained from "foreign" consumers (See equation (3)). Banks set both a deposit fee P_i and variable fees f_i and w_i to attract "home" consumers. Therefore, it is as if the issuers competed in two-part tariffs on the market for deposits.

Competition in two-part tariffs generates pricing at the average perceived cost for the variable part. Consumers internalize the expected transaction costs born at stage 4 when they choose to open an account at stage 3. Hence, a bank can encourage efficient usage of payment instruments at stage 4 by pricing the transactions at the average perceived cost, while extracting surplus from the consumers through the deposit fee.[39] Notice that the card fee takes into account the average cost of cash for the bank, $\varphi_i k\lambda_i^*$, as banks subsidize the consumers by lowering the card fee to encourage them to renounce to cash payments. If the consumers pay all their expenses by card, the card fee is equal to the perceived marginal cost of the Issuer, that is $f_i^*=c_I-a^C$. The withdrawal fee, which is charged only for "foreign" withdrawals, reflects the average cost of withdrawals, as $(1-\varphi_i)w_i^*=(1-\varphi_i)a^W+\varphi_i c_W$.

Each bank extracts a part of the surplus that a consumer obtains from opening an account through the deposit fee. Notice that the transactions made by "home" consumers are subsidized by the deposit fee, as Corollary 1 shows that the transaction margin on home consumers is negative. This analysis explains why banks often argue that payments are loss-leaders.[40] When setting the deposit fee, banks take into account the costs and the benefits of attracting consumers. In my model, two elements soften the competition for deposits. First, the variable costs of cash generated by the consumers who withdraw a lot from their "home" bank reduce the benefits of attracting a consumer. This is reflected by the presence of the terms $-2\left(M_{HC}^{i}\right)^{*}\geq 0$ and $-\left(M_{HC}^{j}\right)^{*}\geq 0$ in the deposit fee. Hence, the variable costs of cash soften the competition for deposits if the proportion of "home" withdrawals is high, as they enable banks to increase the deposit fee. Second, the possibility to make profit on "foreign" consumers lowers the gain of attracting a consumer. Hence, a consumer has to pay the opportunity cost for bank i of losing

[39] It is possible to check that the transaction fees obtained in Proposition 2 maximize $M_i^{HC}-P_i-C_i$, which corresponds to the joint surplus of bank i and of the consumers of bank i on the transaction made by the consumers of bank i.

[40] According to Van Hove (2002), a study conducted by McKinsey on the private costs of Dutch banks finds that in 2005, "the banks in the Netherlands incurred an overall loss of EUR 23 million on their payments business." Van Hove (2002) reports that "Cash generates a loss of no less than EUR 779 million, debit card payments a loss of EUR 101 million (...) That the overall loss is nevertheless limited to EUR 23 million is due to the fact that the income from outstanding balances on retail and corporate accounts is sizeable."

a "foreign" consumer (term $\left(M_{FC}^{i}\right)^{*}$ in the deposit fee). As in Massoud and Bernhardt (2002), or Donze and Dubec (2006), the competition for deposits is softened by the possibility to make profit on "foreign" consumers through the foreign withdrawals.

Finally, I am able to express banks' profit at the equilibrium of stage 2 if the merchants accept cards. For all $(i,j)\epsilon\{1,2\}^2$ and $i \neq j$, bank i's profit at the equilibrium of stage 3 is

$$\pi_i = 2t(\gamma_i^*)^2 + n_j^*(1-\varphi_j)(a^W - c_W) - k(1-\varphi_j)S(\lambda_j^*). \tag{6}$$

In the following section of the paper, I will focus on the case in which banks are perfectly symmetric,[41] and I will denote their joint profit in this case by π, where

$$\pi = t + 2n^*(1-\varphi)(a^W - c_W) - 2k(1-\varphi)S(\lambda^*). \tag{7}$$

3.5 Stage 1: the profit maximising interchange fees in a symmetric equilibrium.

In this section, I study the choice of the interchange fees if banks are perfectly symmetric. I start by analyzing how the level of interchange fees impact the consumers' payment decisions at stage 4 and the competition for deposits. Finally, I determine the interchange fees for card payments and cash withdrawals that maximise banks' joint profits. In all this section, I assume that the function F is chosen such that the second-order conditions of profit maximisation are verified.[42]

3.5.1 Impact of interchange fees on consumers' payment decisions.

The levels of interchange fees impact the consumers' payment decisions and the consumers' transaction costs through the transaction fees that are chosen by the banks at stage 2. Lemma 4 gives the condition under which the consumers use both payment instruments at the equilibrium.

Lemma 4 *If the merchants accept cards and if $a^W \geq c_W$, the consumers never pay cash for all their expenses, and there exists a threshold denoted by $\widehat{a_i^C}(a^W)$ such that the consumers of bank i use both cards and cash if $a^C < \widehat{a_i^C}(a^W)$.*

Proof. See Appendix C-1. ■

[41] Banks are perfectly symmetric if $\varphi_1 = \varphi_2$ and $v_1 = v_2$. Asymmetries between banks will be reintroduced in the extension Section.

[42] This is the case, for instance if $F(T) = 2ST$ over the interval $[0;1]$ and if the costs of cash are small. Indeed, we have that $\partial^2\pi/\ \partial^2 a^C = -2k(1-\varphi)(\ \partial\lambda^*/\ \partial a^C)^2 F'(\lambda^*)$. The determinant of the Hessian matrix is negative if k is sufficiently small, as $(\partial^2\pi/\ \partial^2 a^C)(\partial^2\pi/\ \partial^2 a^W) - (\partial^2\pi/\ \partial a^W \partial a^C)^2 < 0$.

The consumers use both cash and cards to pay for their expenses if the interchange fee on card payments is not too high. Lemma 5 shows that higher interchange fees encourages consumers to substitute cards for cash if β is sufficiently small.

Lemma 5 *If the interchange fee on withdrawals increases, the consumers choose more often to pay by card, and make fewer withdrawals. The total transaction costs born by the consumers increase. When the interchange fee on card payments increases, if β is sufficiently small, the threshold above which consumers pay by card decreases, and the consumers make fewer withdrawals. The total transaction costs born by the consumers decrease.*

Proof. See Appendix C-2. ■

If the consumers use a combination of card and cash payments, the effect of an increase in a^C is to reduce the perceived marginal cost of the issuers for card payments. As the issuers price the transactions at the average cost, the card fee becomes lower. Consumers choose more often to pay by card and make fewer withdrawals, provided that the surcharge rate for card payments is not too high. An increase in a^W raises the average perceived marginal cost of each bank, if consumers make some foreign withdrawals. The withdrawal fee becomes higher and the consumers reduce their volume of cash payments.

3.5.2 The profit maximising interchange fees.

At stage one, the payment platform chooses the interchange fees that maximise banks' joint profit. If there is an interior solution, it verifies the first order conditions of joint profit maximisation, that is, from (7),

$$\frac{\partial \pi}{\partial a^W} = 2n^*(1-\varphi) + 2\frac{\partial n^*}{\partial a^W}(1-\varphi)(a^W - c_W) - 2k(1-\varphi)\frac{\partial \lambda^*}{\partial a^W}F(\lambda^*) = 0, \tag{8}$$

and

$$\frac{\partial \pi}{\partial a^C} = 2\frac{\partial n^*}{\partial a^C}(1-\varphi)(a^W - c_W) - 2k(1-\varphi)\frac{\partial \lambda^*}{\partial a^C}F(\lambda^*) = 0. \tag{9}$$

The following proposition gives the profit maximising interchange fees.

Proposition 3 *If there is an interior solution, the profit maximising interchange fees on withdrawals is higher than the monopoly price, that is,*

$$\frac{a^W - c_W}{a^W} = \frac{1}{\epsilon} + \frac{kF(\lambda^*)}{a^W}\frac{\partial \lambda^*/\partial a^W}{\partial n^*/\partial a^W},$$

where ϵ denotes the elasticity of the number of withdrawals to the interchange fee on withdrawals. The profit maximising interchange fee on card payments is chosen such that the marginal benefits of withdrawals are equal to the marginal costs of withdrawals, that is,

$$(a^W - c_W)\frac{\partial n^*}{\partial a^C} = k\frac{\partial \lambda^*}{\partial a^C}F(\lambda^*).$$

If $k = 0$, the profit maximising interchange fee on card payments is equal to zero, and the profit maximising interchange fee on withdrawals is equal to the monopoly price. If $\varphi = 1$, the level of interchange fees has no impact on banks' joint profit.

If there is an interior solution, the profit maximising interchange fee on card payments reflects the trade-offs that the issuing banks make between the profit made on "foreign" withdrawals, and the possibility to save the variable costs of cash. When banks increase the interchange fee on card payments, this reduces the profit that they make on "foreign" withdrawals, as the consumers pay more often by card (term $(a^W - c_W)\partial n^*/\partial a^C$ in (9)). However, this enables them to reduce the variable costs of cash, as consumers withdraw less cash from the ATM network (term $(a^W - c_W)\partial \lambda^*/\partial a^C$ in (9)). If the variable costs of cash are equal to zero, the profit maximising interchange fee on card payments is optimally equal to zero, as banks can increase the profit that they make on "foreign" withdrawals by encouraging the consumers to renounce to card payments and pay cash. If the variable costs of cash are very high, the profit maximising interchange fee on card payments is chosen such that all consumers pay by card for all their expenses, provided that this interchange fee does not exceed the threshold above which the merchants refuse cards. The profit maximising interchange fee on withdrawals is set by the platform such that the markup is equal to the sum of the inverse of the elasticity of the number of withdrawal to the interchange fee and a positive term, that reflects the variable costs of cash for the consumers.

Let me compare the result of Proposition 3 with the existing literature on payment cards. In my model, the merchants are homogenous as regards to the benefit of accepting cards and there is perfect competition on the acquisition market. In several papers of the literature that make this assumption (see for instance Rochet and Tirole (2002)), the profit maximising interchange fee on card payments is set at the maximum level that is compatible with merchants' acceptance of payment cards. This is not the case in my paper if the variable costs of cash are not too high. By modelling explicitly the costs of cash for the issuing banks, and the profits made on foreign withdrawals, my paper shows that the issuing banks choose an interchange fee that is not necessarily high, as they face themselves a trade-off between the profits made on the ATM side of the market and the profits made on the card side of the market.

3.6 Stage one: the welfare maximising interchange fees in a symmetric equilibrium.

In this section, I start by analysing the impact of interchange fees on consumer and merchant surplus. Then, I determine if the profif maximising interchange fees are too high or too low to maximise the total user surplus. Finally, I study the welfare maximising interchange fees.

3.6.1 Analysis of the impact of interchange fees on merchant and consumer surplus.

Impact of interchange fees on merchant surplus. The merchant surplus, denoted by MS, is the volume of total spending made by the consumers, minus the costs of accepting each payment instrument, that is,

$$MS = S - (1-\beta)(a^C + c_A)(S - S(\lambda^*)) - c_M S(\lambda^*). \tag{10}$$

Merchants accept payment cards as long as it is less costly for them to do so. Hence, if merchants accept cards, the inequality $a^C + c_A \leq c_M/(1-\beta)$ holds and the cost of card payments for merchants is always lower that the cost of cash.[43] A rise in the interchange fee on card payments has two opposite effects on the merchants' surplus. It may increase the volume of transactions that are paid by card (a positive effect), while it decreases the margin that a merchant obtains each time a consumer pays by card (a negative effect). On the contrary, the merchant surplus increases with the interchange fee on withdrawals, as a higher interchange fee on withdrawals increases the volume of transactions that are paid by card.[44] As noted by Van Hove (2002), one way to encourage consumers to pay by card without reducing the merchant surplus is to increase the cost of cash.

Impact of interchange fees on consumer surplus. The consumer surplus, denoted by CS, depends on the total costs that are born by the consumers and on the surplus of opening an account and making purchases, that is

$$CS = V + B - S - (P^* + C^*), \tag{11}$$

[43] In this model, for simplicity reasons, we abstract from modelling the strategic interactions between merchants. According to Rochet and Tirole (2002), merchants may accept cards even if their cost is higher than the cost of cash if there is imperfect competition on the product market. The impact of merchant heterogeneity will be discussed in the extension section.

[44] See Appendix D-1 for the derivatives of the merchant surplus with respect to the interchange fee.

where from Proposition 2 and Corollary 1,

$$P^* = t + \left[n^*(1-\varphi)(a^W - c_W) - k(1-2\varphi)S(\lambda^*) + k\varphi\lambda^* \int_{\lambda^*}^{\bar{\lambda}} \frac{F(T)}{T} dT \right], \tag{12}$$

and from (1),

$$C^* = \sqrt{2r((1-\varphi)a^W + \varphi c_W + b)S(\lambda^*)} + (c_I - a^C - k\varphi\lambda^*) \int_{\lambda^*}^{\bar{\lambda}} \frac{F(T)}{T} dT + (\beta(a^C + c_A) - v)(S - S(\lambda^*)). \tag{13}$$

The impact of interchange fees on consumer surplus is complex. This is because the deposit fee reflects the intensity of competition for deposits, which depends on the trade-offs that banks make between home and foreign consumers. A higher a^W increases the total transaction costs of the consumers and may soften the competition for deposits, if it increases the margin that issuing banks make on "foreign" consumers. A higher a^C decreases the transaction costs born by the consumers and toughens the competition for deposits, if it decreases the margin on "foreign" consumers. In general, it is impossible to sign the variation of the consumer surplus with the interchange fees (See Appendix D). However, one can prove that if k and β are small, and if $\varphi \neq 1$, the consumer surplus increases with a^C.

Profit maximising interchange fees and total user surplus. In Proposition 4, I compare the profit maximising interchange fees with the interchange fees that maximise the total user surplus, which I define as the sum of the consumer surplus and the merchant surplus.

Proposition 4 *The profit maximising interchange fee on card payments is too high to maximise the total user surplus if the cost of cash is low for banks and for merchants, and if the number of transactions paid by card is low compared to the value of card payments. The profit maximising interchange fee on card payments is too low to maximise the total user surplus only if the cost of cash is high for banks and for merchants, and if the value of card payments is low. The profit maximising interchange fee on withdrawals is too high to maximise the total user surplus if the costs of cash are small for banks and merchants, and if the volume of foreign withdrawals is relatively high.*

Proof. See Appendix D-3. ■

The result of Proposition 4 proves that measuring the costs of cash for the agents involved in the payments process (banks, merchants, consumers) is an essential step to assess if the interchange fees are too high or too low from the perspective of social welfare maximisation. If the costs of cash for banks are low, the profit maximising interchange fees on card payments and cash withdrawals are likely to be too high to maximise total user surplus. A high interchange fee on card payments benefits the consumers, but is costly for merchants, especially when the volume of card payments is already high, and when there is no need to encourage consumers to substitute cards for cash. A high interchange fee on withdrawals is not so beneficial for merchants if the benefit of accepting cards is low, and it is detrimental to the consumers as it increases the costs of foreign withdrawals, which cannot be avoided by the consumers if some of them are captive from the banks.

The only case in which the interchange fee on card payments is too low to maximise total user surplus is the situation in which the cost of cash is high for banks or for merchants, and the value of card payments is low. This situation is also detrimental to the consumers, as banks pass through the costs of cash to consumers through the deposit fee. One policy implication of this result would be to encourage banks to use interchange fees on card payments in the countries where the use of payment cards is too low, and when it could enable banks to cut the cost of cash.

3.6.2 Welfare analysis.

I now study the impact of interchange fees on social welfare. A benevolent social planner chooses the interchange fees a^C and a^W that maximise social welfare, which is the sum of banks profit, consumer surplus and merchant surplus, that is $W = \pi + MS + CS$. The derivatives of W with respect to a^C and a^W are given in Appendix D-4.

The welfare maximising interchange fees balance the interests of the banks, the consumers and the merchants. I do not determine the welfare maximising interchange fees as the computations are complex. However, I provide here some intuitions about their determination. If the benefit of accepting cards for merchants is high (for instance if c_M is high), and if the costs of cash for banks are high, it is socially optimal to choose a high interchange fee on cash withdrawals, that will discourage consumers to pay cash (as term C in (39) of Appendix D-4 is higher in this case than terms A and B). If the benefit of accepting cards for merchants is low (term C in (39)), it may be socially optimal to have a low interchange fee on withdrawals, provided that the costs of cash for banks and for consumers (term B in (39)) are not too high, and if consumers make a lot of foreign withdrawals (φ small in term A).

The issuers compete on the market for deposits, which enables them to internalize a part of the consumer's surplus when they choose the interchange fee on card payments. However, they have an incentive to choose a lower interchange fee on card payments than the one that would be optimal for consumers, because they can make profit on foreign withdrawals (terms E and F are positive in (40)). This level of interchange fee may be still to high to maximise merchant surplus if the value of the transactions paid by card is high (term G), and if the cost of cash for merchants is low (term H). The fact that the issuers' interests are not aligned with the interests of the joint user (consumer+merchant) is another source of inefficiency in the substitution between cash and cards.

4 Extensions and discussions.

In this section, I provide some extensions of the results obtained in section 3. First, I discuss how Proposition 3 and 4 would be impacted by merchants' heterogeneity. Second, I consider the case of asymmetric issuers. Then, I study the case in which symmetric issuers are also acquirers. Finally, I discuss the case of endogeneous ATM deployment decisions.

4.1 Heterogeneous merchants.

In this subsection, I assume that the merchants are heterogeneous over the cost of accepting cash payments c_M, such that a share α of the merchants accept cards and cash, while a share $1-\alpha$ of the merchants accept only cash. The share of the merchants who accept cards is decreasing with the merchant fee and with the interchange fee on card payments, as $m = a^C + c_A$. When the consumers choose how often to withdraw from the ATM network, they anticipate that the debit card has a probability α of being accepted for a transaction, and a probability $1-\alpha$ of being refused. In Appendix F-1, I prove that the number of withdrawals and the threshold λ decrease with the share of the merchants who accept cards. At stage 2, the issuing banks choose the same transaction fees as in Proposition 2, but the deposit fee increases with the percentage of merchants who accept cards. Banks' joint profit now depends on the share of merchants who accept cards, that is

$$\pi = t + 2n^*(1-\varphi)(a^W - c_W) - 2k(1-\varphi)\left[\alpha S(\lambda^*) + (1-\alpha)S\right].$$

When the share of merchants who accept cards increases, the cost of cash for the issuing banks is reduced. In Proposition 3, I argued that issuing banks choose interchange fees that reflect a trade-off between the profits on the ATM transactions and the costs of cash. A high interchange fee on card payments reduces the cost of cash as it encourages the consumers to pay by card. However, if merchants' heterogeneity is taken into account, a high interchange fee on card payments also reduces merchants' acceptance of payment cards, which *increases* the costs of cash for the issuing banks, as consumers pay less by card. Hence, the trade-off between the profits on ATM transactions and the costs of cash is impacted by the assumption that merchants are heterogeneous. Proposition 5 gives the profit maximising interchange fees under merchants' heterogeneity.

Proposition 5 *If merchants are heterogeneous, and if there is an interior solution, the profit maximising interchange fees verify*

$$\frac{a^W - c_W}{a^W} = \frac{1}{\epsilon} + \frac{k\alpha^* F(\lambda^*)}{a^W} \frac{\partial \lambda^* / \partial a^W}{\partial n^* / \partial a^W},$$

and

$$(a^W - c_W) \frac{\partial n^*}{\partial a^C} = k\alpha^* \frac{\partial \lambda^*}{\partial a^C} F(\lambda^*) - k \frac{\partial \alpha^*}{\partial a^C} (S - S(\lambda^*)).$$

Proof. See Appendix F-1. ■

Finally, to analyse the effect of interchange fees on consumer and merchant surplus, and to check the robustness of Proposition 4, we have to take into account the fact that a higher interchange fee on card payments reduces the share of merchants who accept cards. This effect is not necessarily detrimental to consumers, as they have the possibility to ajust their payments decisions according to the costs of the payment instruments and the probability that cards are accepted. If the consumers' transaction costs do not increase too much when card acceptance is reduced, the results of Proposition 4 are impacted as follows by merchants' heterogeneity. Proposition 4 remains logically unchanged if merchants acceptance of payment cards is high (See Appendix F-1). However, if merchants' acceptance of payment cards is low, the interchange fee on card payments may be too low to maximise total user surplus. Also, the second part of Proposition 4 could be restated as follows: the profit maximising interchange fee on withdrawals is too high to maximise the total user surplus if the costs of cash are small for banks, if merchants' acceptance of payment cards is limited and if the volume of foreign withdrawals is high.

4.2 Asymmetries between issuers.

The model that is presented in section 3 could be analysed as a situation in which there are two platforms (the ATM network and the debit card platform), controlled by the same group of issuers, which compete to offer substituable payment and withdrawal services to the consumers. In this case, there is symmetric "duality", in the sense that symmetric issuing banks are members of both platforms, which enables them to provide the same services to the consumers at the equilibrium of the game. Under symmetric duality, the issuers' interests are aligned, and they are able to choose interchange fees to extract as much surplus as possible from the consumers, by internalizing the effects of the competition between payment and withdrawal services. The situation could be different if some issuers are dual members (card issuers and ATM owners), whereas other issuers do not own ATMs, or if the issuers are asymmetric (asymmetric dual members).

To understand better the impact of asymmetries between issuers, I study an extreme case in which the first issuer controls all the ATM network, that is $\varphi_1 = 1$, while the second issuer does not own any ATM at all, that is $\varphi_2 = 0$. The consumers of bank 1 make all their withdrawals in their bank for free, while the consumers of bank 2 make only foreign withdrawals, priced at a^W. In this case, from (6), banks' joint profit is

$$\pi = 2t(\gamma_1^*)^2 + 2t(1-\gamma_1^*)^2 + n_2^*(a^W - c_W) - kS(\lambda_2^*),$$

where γ_1^* denotes the market share of the first issuer. As banks are not symmetric, the profit maximising interchange fees do not only reflect bank 1's trade-off between the profit on foreign withdrawals and the costs of cash, but also the impact of interchange fees on banks' market shares. In this case, the issuers' interests are not aligned (See Appendix F-2). If the costs of cash are not too high, bank 1 (the ATM owner) benefits from a high interchange fee on withdrawals, as this increases its market share to the detriments of bank 2 (the pure issuer). Studying banks' or entrants' decisions to become members of ATM and/or debit card platforms is beyond the scope of this paper, but this issue should deserve further investigation, as market shares have an impact on the choice of interchange fees.

4.3 Symmetric issuers as ATM owners and acquirers.

In this subsection, I discuss how my results could change if symmetric issuers compete also on the merchant side. To simplify the model, I assume that β is close to zero. When they make their affiliation decision, the merchants take into account the expected transaction costs at stage 4, denoted by $\left(C_i^M\right)^*$, the fixed deposit fee M_i, and the transportation cost t_M, which depends on their location. A merchant located at point $x \in [0;1]$, that opens an account at bank i located at d_i, bears a cost $t_M\left|x-d_i\right| + M_i + \left(C_i^M\right)^*$, and obtains a surplus, which I assume to be sufficiently large such that the market is covered. The marginal merchant is given by:

$$\alpha_1 = \frac{1}{2} + \frac{1}{2t_M}\left[M_2 - M_1 + \left(C_2^M\right)^* - \left(C_1^M\right)^*\right].$$

When a consumer of bank 1 pays by card, I assume that he has a probability α_1 to shop at a merchant's who is affiliated at bank 1 and $1-\alpha_1$ to shop at a merchant's who is affiliated at bank 2. In Appendix F-3, I prove that banks' profit in a symmetric equilibrium is

$$\pi = \frac{t_C}{2} + \frac{t_M}{2} + n^*(1-\varphi)(a^W - c_W) - k(1-\varphi)S(\lambda^*),$$

where t_C denotes the transportation cost of the consumers, when they travel to open an account. As the variable part of π is exactly the same as in my main model, the profit maximising interchange fees remain identical to Proposition 3, provided that the issuers-acquirers are symmetric. If the issuers-acquirers are not symmetric, one could suspect that the issuing bank that has the lowest market share on the acquisition side would benefit from higher interchange fees on card payments and lower interchange fees on withdrawals, all other things being equal.

4.4 Endogenous ATM deployment decisions.

In the model of section 3, I assumed that the percentage of foreign withdrawals made by the consumers was exogenous. I examine in this subsection how the results would change if ATM deployment decisions are endogenous. I assume that φ_1 represents now the probability that a consumer of bank 1 withdraws cash from an ATM owned by bank 1. The probability of a "home" withdrawal is related to banks' deployment decisions as follows: if ρ_i denotes the number of ATM of bank i for all $i \epsilon \{1, 2\}$, the probability that a consumer of bank i withdraws cash from his bank is $\varphi_i = \rho_i/(\rho_1 + \rho_2)$. Let $DC(\rho)$ be the cost of deploying ρ ATMs, and assume that banks' costs functions are identical and convex. After the choice of interchange fees, banks have to decide how many ATMs to deploy (this stage would be added after the first stage of the initial game presented in section 3).

In Appendix F-4, I prove that the number of withdrawals made by a consumer and the threshold λ increase with the number of ATMs deployed by his bank. The intuition of this result is that an increase in the number of ATMs owned by his "home" bank increases the probability for a consumer to make a free withdrawal, which encourages him to withdraw cash more often, and pay less by card. In a symmetric equilibrium, banks deploy the same number of ATMs, such that the marginal benefit of investments in ATMs is equal to the marginal cost, that is

$$\frac{1}{2}\frac{\partial n^*}{\partial \rho}(a^W - c_W) + \frac{n^*}{2\rho}(a^W - c_W) = k\frac{\partial \lambda^*}{\partial \rho}F(\lambda^*) + \frac{k}{2\rho}S(\lambda^*) + DC'(\rho). \tag{14}$$

The marginal benefit of investing in ATM deployment is linked to the possibility to make profits on foreign withdrawals, as in Donze and Dubec (2006). However, in my setting, banks take also into account the costs of having to manage higher volumes of cash in their investment decision (term $k\dfrac{\partial \lambda^*}{\partial \rho}F(\lambda^*)$ in (14)). At stage 1, the profit maximising interchange fees verify exactly the same equations as in Proposition 3 in a symmetric equilibrium, except that $\varphi = 1/2$, as banks deploy the same number of ATMs when their deployment costs are identical. Considering endogenous ATM deployment decisions is useful to study a different issue, which is not the initial purpose of this paper: is there an excessive deployment of ATMs due to interchange fees? My model adds another ingredient to the framework developped by Donze and Dubec (2006). As in Donze and Dubec (2006), a higher interchange fee on withdrawals increases the incentives to invest in ATM deployment, as it increases the marginal benefit that banks can earn from foreign withdrawals (if second-order effects are neglected). However, unlike Donze and Dubec (2006), the ATM deployment is slowed down by higher interchange fees on card payments, as they reduce the number of foreign withdrawals, and consequently the marginal benefits of deploying ATMs.

5 Conclusion and discussion.

In this article, I have explained why the collective choice of interchange fees for debit cards leads to inefficiencies in the substitution between cash and card payments, when the issuers are also ATM owners. A profit-maximising payment platform chooses interchange fees that reflect banks' trade-off between the revenues on foreign consumers, the revenues on deposits, and the costs of cash. From a social welfare point of view, the interchange fees on card payments are too high if the cost of cash for banks and merchants is low, but may be too low if the cost of cash either for banks or merchants is high.

This analysis has led me to argue that measures of the costs of cash are essential to assess if the interchange fees are too high or too low. I have also suggested that banks can react by levying higher fees on the ATM side of the market in response of a regulation of interchange fees on card payments. Interestingly, Chakravorti, Carbo-Valverde and Rodriguez (2009) note that "surcharges for foreign ATM withdrawals have been increasing for Spain", during a period in which interchange fees on card payments have been regulated. One could suspect that issuing banks have used surcharges to recover interchange fees losses on the card side, but this should deserve further investigation.

Inefficiencies in the substitution between cash and cards have already been analysed in several surveys conducted by Central Banks (See for instance Brits and Winder (2005) for the Netherlands). Following these studies, some political remedies have been considered to lower the switching point between cash and card payments. For instance, the DNB has measured the impact of an increase in the number of POS terminals and the impact of a halt in the increase of the number of ATMs.[45] Other authors (See Van Hove (2004)) have argued that the introduction of cost-based pricing for payments would be the best solution: "rather than concentrating on the introduction of charges for services that are currently cross-subsidized, policy makers might also try to remove the sources of this cross-subsidization". A policy that could be considered, which is not introduced in my model, is the promotion of entry in payments markets. Entry has been encouraged for instance in Australia, where consumers have to pay a usage fee to the owner of the ATM, according to the "direct charging reform". Multilateral interchange fees on ATMs have been suppressed in Australia, whereas interchange fees on card payments have been capped by the regulator for Visa and MasterCard.[46] Further research is needed to understand whether the promotion of competition, either between banks and non-banks, or between platforms, is the best way to remove the inefficiencies in the substitution between payment cards and cash.

References

[1] APACS, The way we pay 2008, "UK Cash & Cash Machines".

[2] ARANGO, C. & TAYLOR, V. (2008): "Merchants Acceptance, Costs and Perception of Retail Payments: A Canadian Survey," Bank of Canada, Discussion Paper 2008-12.

[3] BAXTER W. (1983): "Bank Interchange of Transactional Paper: Legal and Economic Perspectives," Journal of Law & Economics, vol. 26, no. 3, October, pp. 541-588.

[4] BAUMOL W. (1952): "The Transactions Demand for Cash: An Inventory Theoretic Approach," The Quaterly Journal of Economics, vol. 66, no. 4, November, pp. 545-556.

[5] BERGMAN M., GUIBOURG G. & SEGENDORF B. (2007): "The Costs of Paying - Private and Social Costs of Cash and Card Payments," Riksbank Research Paper Series N.212.

[6] BRITS H. & WINDER C. (2005): "Payments Are No Free Lunch," De Nederlandsche Bank, Occasional Studies, vol. 3 no.2.

[7] BOESCHOTEN W. (1998): "Cash Management, Payment Patterns and the Demand for Money," De Economist, 146(1): 117-142.

[45] See DNB Quaterly Bulletin, March 2006.

[46] See: http://www.rba.gov.au/payments-system/reforms/atm/access-regime/reform-package.html. "In particular, the move to a regime in which the ATM owners directly charge cardholders rather than earn revenue through interchange fees will increase the competition for the provision of ATM services (...). In addition, the access reforms will make it easier for new firms to enter the market".

[8] BOUNIE D. & FRANCOIS A. (2006): "Cash, Check or Bank Card? The effect of transaction characteristics on the use of payment instruments," mimeo.

[9] BORZEKOWSKI R., KISER E., AHMED S. (2008): "Consumers' Use of Debit Cards: Patterns, Preferences, and Price Response," Journal of Money, Credit and Banking, vol. 40(1), pp.149-172.

[10] BERNHARDT D. & MASSOUD N. (2002): ""Rip-off" ATM surcharges," The RAND Journal of Economics, vol. 33, Spring, no. 1, pp. 96-115.

[11] BOLT W. & CHAKRAVORTI S. (2008): "Consumer Choice and Merchant Acceptance of Payment Media," DNB Working Paper no.197.

[12] CHAKRAVORTI S., CARBO-VALVERDE & RODRIGUEZ FERNANDEZ (2009):"The Costs and Benefits of Interchange Fee Regulation: An Empirical Investigation," Working Paper.

[13] CHAKRAVORTI S. & ROSON R. (2006): "Platform Competition in Two-Sided Markets: The Case of Payment Networks," Review of Network Economics, vol. 5(1), March, pp. 118-143.

[14] CHAKRAVORTI S. & TO T. (2007): "A Theory of Credit Cards," International Journal of Industrial Organization, vol 25(3), pp. 583-595.

[15] CHANG H., EVANS D. S., & SWARTZ D. GARCIA (2005): "The Effect of the Regulatory Intervention in Two-Sided Markets: An Assessment of Interchange-Fee Capping in Australia," Review of Network Economics, vol. 4, no. 4, December, pp. 328-358.

[16] DE GRAUWE, P., RINALDI, L. & VAN CAYSEELE, P. (2006): "Issues of efficiency in the use of cash and cards," University of Louvain, Mimeo.

[17] De Nederlandsche Bank, "Towards a Cashless Society?," Quaterly Bulletin, March 2006, pp.63-70.

[18] DONZE J. & DUBEC I. (2006): "The Role of Interchange Fees in ATM networks," International Journal of Industrial Organization, vol. 24, January, pp. 29-43.

[19] DONZE J. & DUBEC I. (2009): "Pay for ATM Usage: Good for Consumers, Bad for Banks?," Journal of Industrial Economics, vol. 59, September, pp. 583-612.

[20] GUTHRIE G. & WRIGHT J. (2007): "Competing Payment Schemes," Journal of Industrial Economics, vol. 55, pp. 37-67.

[21] LAFFONT J-J., REY P., TIROLE J. (1998): "Network competition: Overview and nondiscriminatory pricing," RAND Journal of Economics, vol. 29, Spring 1998, pp. 1-37.

[22] McANDREWS J. (2003) "Automated Teller Machine Network Pricing - A Review of the Literature," Review of Network Economics, vol. 2, Issue 2, June, pp. 146-158.

[23] MATUTES C. & PADILLA A.J (1994): "Shared ATM networks and Banking Competition," European Economic Review, vol. 38, pp. 1113-1138.

[24] RAA ten, T. & SHESTALOVA, V. (2004): "Empirical evidence on payment media costs and switch points," Journal of Banking & Finance, Elsevier, vol. 28, Issue 1, January, pp. 203-213.

[25] ROCHET J-C. (2003): "The Theory of Interchange Fees: A synthesis of recent contributions," Review of Network Economics, vol. 2, no. 2, June, pp. 97-124.

[26] ROCHET J-C. & TIROLE J. (2002): "Cooperation Among Competitors: The Economics of Payment Card Associations," The RAND Journal of Economics, vol. 33, no. 4, winter, pp. 549-570.

[27] ROCHET J-C & TIROLE (2003): "Platform Competition in Two-Sided Markets," Journal of the European Economic Association, vol. 1, no. 4, pp. 990-1029.

[28] ROCHET J-C. & TIROLE J. (2006): "Two-sided markets: a progress report," RAND Journal of Economics, vol. 37, no. 3, March, pp. 645-667.

[29] ROCHET J-C. & TIROLE J. (2009): "Must-Take Cards: Merchant Discounts and Avoided Costs," Journal of the European Economic Association (Forthcoming).

[30] SCHOLNICK B., MASSOUD N., SAUNDERS A., CARBO-VALVERDE S. & RODRIGUEZ-FERNANDEZ F. (2008): "The Economics of Credit Cards, Debit Cards and ATMs: A Survey and Some New Evidence," Journal of Banking and Finance, 32(8), pp.1468-1483.

[31] SCHUH S. & STAVINS J. (2009): "Why Are (Some) Consumers (Finally) Writing Fewer Checks? The Role of Payment Characteristics," Journal of Banking and Finance (Forthcoming).

[32] SHY O. & WANG Z. (2010): "Why do Card Issuers Charge Proportional Fees", American Economic Review (Forthcoming).

[33] TOBIN J. (1956): "The Interest-Elasticity of Transactions Demand For Cash," The Review of Economics and Statistics, vol. 38, no. 3 (August), pp. 241-247.

[34] VAN HOVE L. (2004): "Cost-based pricing of payment instruments: the state of the debate," De Economist, vol. 152, no.1, pp.79-100.

[35] VERDIER M. (2010): "Interchange fees and incentives to invest in quality of a payment card system," International Journal of Industrial Organization (Forthcoming).

[36] VERDIER M. (2010): "Interchange fees in Payment Card Systems: a Survey," Journal of Economic Surveys (Forthcoming).

[37] WANG Z. (2008): "Regulating Debit Cards: the Case of Ad Valorem Fees", Federal Reserve Bank of Kansas City.

[38] WHITESELL (1989): "The Demand for Currency versus Debitable Account: Note," Journal of Money, Credit and Banking, vol. 21, no. 2 (May), pp. 246-251.

[39] WRIGHT J. (2002): "Optimal Payment Card Systems," European Economic Review, vol. 47, no. 4, August, pp. 587-612.

[40] WRIGHT J. (2004): "Determinants of Optimal Interchange Fees in Payment Systems," Journal of Industrial Economics, vol. 52, no. 1, March, pp. 1-26.

6 Appendix

Appendix A: Proof of Proposition 1 and some comparative statics

Appendix A-1: Proof of Proposition 1. The consumer chooses the threshold λ_i and the number of withdrawals n_i that minimize its costs. I start by determining the equations verified by an interior optimum, if $\lambda_i > 0$ and $n_i > 0$. Then I derive the conditions that are necessary and sufficient for this optimum to exist. Solving for the first order conditions of profit maximisation[47], I obtain

$$\frac{\partial C_i(\lambda_i, n_i)}{\partial n_i} = (1 - \varphi_i)w_i + b - \frac{rS(\lambda_i)}{2(n_i)^2} = 0, \tag{15}$$

and

$$\frac{\partial C_i(\lambda_i, n_i)}{\partial \lambda_i} = \frac{r}{2n_i}\frac{\partial S(\lambda_i)}{\partial \lambda_i} - f_i\frac{F(\lambda_i)}{\lambda_i} + (v_i - \beta m)\frac{\partial S(\lambda_i)}{\partial \lambda_i} = 0. \tag{16}$$

[47] The Hessian matrix is semi definite negative as in Whitesell (1989).

Let us denote an interior solution by (λ_i^*, n_i^*), where λ_i^* is the threshold above which consumers pay by card, and n_i^* the optimal number of withdrawals. From (15), I obtain the optimal number of withdrawals,

$$n_i^* = \sqrt{\frac{rS(\lambda_i^*)}{2((1-\varphi_i)w_i + b)}}. \tag{17}$$

As $S(\lambda_i) = \int\limits_0^{\lambda_i} F(T)dT$, we have $\partial S(\lambda_i)/\partial \lambda_i = F(\lambda_i)$. Replacing for n_i^* in (16), if there is an interior solution, λ_i^*, then it must satisfy

$$F(\lambda_i^*)\left[\sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda_i^*)}} - \frac{f_i}{\lambda_i^*} + (v_i - \beta m)\right] = 0, \tag{18}$$

that is, as $\lambda_i^* > 0$,

$$\lambda_i^* \sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda_i^*)}} - f_i + (v_i - \beta m)\lambda_i^* = 0. \tag{19}$$

I now show that, under some conditions, there exists a unique λ_i^* that verifies equation (19). For this purpose, let

$$g(\lambda) = \lambda \sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda)}} - f_i + \lambda(v_i - \beta m).$$

My aim is to derive the conditions under which there exists a unique λ_i^* such that $g(\lambda_i^*) = 0$. The function g is differentiable over $\left(0; \overline{\lambda}\right]$, and its first derivative is

$$g^{'}(\lambda) = (v_i - \beta m) + \sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda)}}\left(1 - \frac{\lambda F(\lambda)}{2S(\lambda)}\right).$$

Let the function h be defined as $h(\lambda) = 2S(\lambda) - \lambda F(\lambda)$ over $\left[0, \overline{\lambda}\right]$. We have

$$g^{'}(\lambda) = (v_i - \beta m) + \sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda)}}\frac{h(\lambda)}{2S(\lambda)}.$$

I now study the function h over $\left[0, \overline{\lambda}\right]$. As $S^{'}(\lambda) = F(\lambda)$, then $h^{'}(\lambda) = F(\lambda) - \lambda F^{'}(\lambda)$, and $h^{"}(\lambda) = -\lambda F^{"}(\lambda)$. As F is concave by assumption (A2), h is convex. So $h^{'}$ is increasing over $\left[0, \overline{\lambda}\right]$. As $h^{'}(0) = F(0) \geq 0$, $h^{'}$ is positive over $\left[0, \overline{\lambda}\right]$. Hence, h is increasing over $\left[0; \overline{\lambda}\right]$. As $h(0) = 0$, h is positive over $\left[0, \overline{\lambda}\right]$. Hence, as $v_i - \beta m > 0$ if β is small[48], the function $g^{'}$ is strictly positive over $\left(0, \overline{\lambda}\right]$ and consequently, g is increasing over $\left(0, \overline{\lambda}\right]$. By assumption (A3),

[48] This condition is satisfied in equilibrium because of Assumption (A6).

we have

$$\lim_{\lambda \longrightarrow 0}(\lambda/\sqrt{S(\lambda)}) = l.$$

Hence g can be prolongated to $g(0) = l\sqrt{r((1-\varphi_i)w_i + b)/2} - f_i$. If $f_i > l\sqrt{r((1-\varphi_i)w_i + b)/2}$ and $g(\overline{\lambda}) > 0$, using the bijection theorem, there exists a unique $\lambda_i^* \in \left(0, \overline{\lambda}\right)$ such that $g(\lambda_i^*) = 0$. If $f_i < l\sqrt{r((1-\varphi_i)w_i + b)/2}$ or $g(\overline{\lambda}) < 0$, the equation $g(\lambda) = 0$ does not admit any solution over $\left[0, \overline{\lambda}\right]$. The condition $g(\overline{\lambda}) > 0$ is equivalent to $f_i < (v_i - \beta m)\overline{\lambda} + \overline{\lambda}\sqrt{r((1-\varphi_i)w_i + b)/2S(\overline{\lambda})}$. It can be interpreted as follows. The card fee must be lower than the average cost of cash if the consumer decides to pay everything cash. The average cost of cash comprises the opportunity cost of renouncing to the variable net benefit $v_i - \beta m$ and the opportunity cost of cash detention.

If $f_i \leq l\sqrt{r((1-\varphi_i)w_i + b)/2}$, from (16) the consumer's total cost increases with λ_i as g is positive over $\left[0, \overline{\lambda}\right]$. Hence, the optimal threshold λ_i^* is equal to zero, and the consumer pays by card all his expenses. If the card fee is higher than the average cost of paying everything cash, that is if $g(\overline{\lambda}) < 0$, the consumer does not pay by card. In this case, the optimal threshold is $\overline{\lambda}$, and the number of withdrawals is $n_i^* = \sqrt{rS/2((1-\varphi_i)w_i + b)}$.

Appendix A-2: an example with the function $F(T) = 2ST$ **with** $T \in [0; 1]$**.** In this example, we have $S(\lambda) = S\lambda^2$ and $l = \lim\limits_{\lambda \to 0} \dfrac{\lambda}{\sqrt{S\lambda^2}} = \dfrac{1}{\sqrt{S}}$. Hence, from (17) and (19), if $f_i > \sqrt{r((1-\varphi_i)w_i + b)/2S}$, we have

$$n_i^* = \lambda_i^* \sqrt{\frac{rS}{2((1-\varphi_i)w_i + b)}},$$

and

$$\lambda_i^* = \frac{1}{(v_i - \beta m)}\left(f_i - \sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S}}\right).$$

If $f_i \leq \sqrt{r((1-\varphi_i)w_i + b)/2S}$, we have $\lambda_i^* = 0$ and $n_i^* = 0$.

Appendix A-3: Proof of Lemma 1. I now show that λ_i^* and n_i^* increase with f_i and m, and that they decrease with w_i and v_i. I start by showing that λ_i^* and n_i^* increase with f_i. Taking the derivative of (19) with respect to f_i, I obtain that

$$-1 + (v_i - \beta m)\frac{\partial \lambda_i^*}{\partial f_i} + \frac{\partial \lambda_i^*}{\partial f_i}\sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda_i^*)}}\left(1 - \frac{F(\lambda_i^*)\lambda_i^*}{2S(\lambda_i^*)}\right) = 0,$$

that is, after some rearrangements,

$$\frac{\partial \lambda_i^*}{\partial f_i}\left(v_i - \beta m + \sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda_i^*)}}\frac{h(\lambda_i^*)}{2S(\lambda_i^*)}\right) = 1, \tag{20}$$

where $h(\lambda) = 2S(\lambda) - \lambda F(\lambda)$. I have already shown in Appendix A-1 that h is positive over $\left[0, \overline{\lambda}\right]$. Therefore, all the terms in the parenthesis of (20) are positive if $v_i - \beta m > 0$, and I can conclude that $\partial \lambda_i^*/\partial f_i \geq 0$ if β is sufficiently small. Taking the derivative of (17) with respect to f_i, I find that

$$\frac{\partial n_i^*}{\partial f_i} = \frac{F(\lambda_i^*)n_i^*}{2S(\lambda_i^*)}\frac{\partial \lambda_i^*}{\partial f_i}. \tag{21}$$

As $\partial \lambda_i^*/\partial f_i \geq 0$, the number of withdrawals increases with f_i.

Then I show that λ_i^* and n_i^* decrease with w_i. Taking the derivative of (19) with respect to w_i, I obtain that

$$\frac{\partial \lambda_i^*}{\partial w_i}\left(v_i - \beta m + \sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda_i^*)}}\left(1 - \frac{F(\lambda_i^*)\lambda_i^*}{2S(\lambda_i^*)}\right)\right) = -\lambda_i^*\sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda_i^*)}}\frac{1}{2((1-\varphi_i)w_i + b)}.$$

The expression in the parenthesis is positive if β is sufficiently small, as in (20). The right side of the equation is negative. So, $\partial \lambda_i^*/\partial w_i \leq 0$. If the price of foreign withdrawals rises, this reduces the threshold above which consumers pay by card.

Taking the derivative of (17) with respect to w_i, I obtain that

$$\frac{\partial n_i^*}{\partial w_i} = \frac{F(\lambda_i^*)n_i^*}{2S(\lambda_i^*)}\frac{\partial \lambda_i^*}{\partial w_i} - \frac{(1-\varphi_i)n_i^*}{2((1-\varphi_i)w_i + b)}, \tag{22}$$

As $\partial \lambda_i^*/\partial w_i \leq 0$, it follows that $\partial n_i^*/\partial w_i \leq 0$. If the price of foreign withdrawals rises, the number of withdrawals decreases unambiguously.

I now show that λ_i^* and n_i^* decrease with v_i. Taking the derivative of (19) with respect to v_i, I obtain

$$-\lambda_i^* + (v_i - \beta m)\frac{\partial \lambda_i^*}{\partial v_i} + \frac{\partial \lambda_i^*}{\partial v_i}\sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda_i^*)}}\left(1 - \frac{F(\lambda_i^*)\lambda_i^*}{2S(\lambda_i^*)}\right) = 0,$$

that is, after some rearrangements,

$$\frac{\partial \lambda_i^*}{\partial v_i}\left(v_i - \beta m + \sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda_i^*)}}\frac{h(\lambda_i^*)}{2S(\lambda_i^*)}\right) = -\lambda_i^*.$$

Using the same reasoning as in Appendix A-2, I obtain that $\partial \lambda_i^*/\partial v_i \leq 0$ and that $\partial n_i^*/\partial v_i \leq 0$ if β is sufficiently small.

Finally, I show that λ_i^* and n_i^* increase with m. We have that

$$\frac{\partial \lambda_i^*}{\partial m}\left(v_i - \beta m + \sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S(\lambda_i^*)}}\frac{h(\lambda_i^*)}{2S(\lambda_i^*)}\right) = \beta\lambda_i^*.$$

Hence, $\partial\lambda_i^*/\partial m \geq 0$, and $\partial n_i^*/\partial m \geq 0$.

Appendix A-4: Proof of Lemma 2. Using the envelop's theorem, I find that

$$\left.\frac{dC_i(\lambda_i, n_i, w_i)}{dw_i}\right|_{(\lambda_i^*, n_i^*)} = \left.\frac{\partial C_i(\lambda_i, n_i)}{\partial w_i}\right|_{(\lambda_i^*, n_i^*)},$$

and that

$$\frac{\partial C_i^*(\lambda_i^*)}{\partial w_i} = (1-\varphi_i)n_i^* \geq 0. \tag{23}$$

Using the same reasoning, I obtain that

$$\frac{\partial C_i^*(\lambda_i^*)}{\partial f_i} = \int_{\lambda_i^*}^{\overline{\lambda}} \frac{F(T)}{T} dT \leq 0. \tag{24}$$

Similarly, I have that

$$\frac{\partial C_i^*(\lambda_i^*)}{\partial v_i} = -(S - S(\lambda_i^*)) \leq 0, \tag{25}$$

and

$$\frac{\partial C_i^*(\lambda_i^*)}{\partial m} = \beta(S - S(\lambda_i^*)) \geq 0. \tag{26}$$

So the consumer's total cost increases with the withdrawal fee, the card fee and the merchant fee, while it decreases with the variable benefit of paying by card.

Appendix B-5: comparison with the joint cost minimizing switching point. The total cost born by the users, which I denote by C_{TU}, is defined as the sum of the consumer's costs and the merchant's costs:

$$\begin{aligned} C_{TU} &= \frac{r}{2n_i}S(\lambda_i) + n_i((1-\varphi_i)w_i + b) + f_i\int_{\lambda_i}^{\overline{\lambda}} \frac{F(T)}{T} dT + (\beta m - v_i)(S - S(\lambda_i)) + \\ &\quad c_M S(\lambda_i) + (1-\beta)m(S - S(\lambda_i)). \end{aligned}$$

Solving for the first order condition of "joint user" cost minimization, we find that

$$(n_i^*)_{TU} = \sqrt{\frac{rS((\lambda_i^*)_{TU})}{2((1-\varphi_i)w_i + b)}},$$

and that the threshold above which the "joint user" would like to pay by card verifies

$$F((\lambda_i^*)_{TU})\left[\frac{r}{2(n_i^*)_{TU}} - \frac{f_i}{(\lambda_i^*)_{TU}} + (v_i - m + c_M)\right] = 0.$$

For a given threshold, the number of withdrawals chosen by the joint user is the same as the number of withdrawals chosen by the consumer. To avoid confusion, let us denote by χ the threshold chosen by the consumer. We have

$$\left.\frac{\partial C_{TU}}{\partial \lambda}\right|_{\lambda=\chi} = \chi(c_M - (1-\beta)m),$$

and

$$\left.\frac{\partial C_{TU}}{\partial \lambda}\right|_{\lambda=(\lambda_i^*)_{TU}} = 0.$$

As C_{TU} is convex in λ for all $m \leq c_M/1-\beta$, we have $\chi \geq (\lambda_i^*)_{TU}$.

Appendix B: Proof of proposition 2. I study the existence of an equilibrium in which the market shares are strictly positive. I determine the candidate equilibrium by solving the first order conditions of profit maximisation. Then, I will verify the second order conditions. In my paper, I chose to focus on the case in which consumers use a combination of cash and card payments. Hence, I will have to provide ex post the conditions under which this is verified at the equilibrium.

Appendix B-1: The first order conditions for profit maximization The function π_i is twice differentiable over $\left(0;\overline{\lambda}\right)$. To simplify the computations, I write

$$\pi_i = \gamma_i A_i(P_i; f_i; w_i) + n_j^*(1-\varphi_j)(a^W - c_W) - k(1-\varphi_j)S(\lambda_j^*),$$

where

$$\begin{aligned} A_i(P_i; f_i; w_i) &= P_i + (f_i + a^C - c_I)\int_{\lambda_i^*}^{\overline{\lambda}} \frac{F(T)}{T} dT + n_i^*((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W) \\ &\quad - n_j^*(1-\varphi_j)(a^W - c_W) - k(\varphi_i S(\lambda_i^*) - (1-\varphi_j)S(\lambda_j^*)) \\ &= P_i + M_i^{HC} - M_j^{FC} \end{aligned}$$

First, notice that n_j^* and λ_j^*, are independent of f_i, w_i and P_i. Hence, solving for the first order conditions of profit maximisation with respect to P_i, f_i and w_i yields

$$\frac{\partial \pi_i}{\partial P_i} = \frac{\partial \gamma_i}{\partial P_i} A_i(P_i; f_i; w_i) + \gamma_i \frac{\partial A_i(P_i; f_i; w_i)}{\partial P_i} = 0, \tag{27}$$

$$\frac{\partial \pi_i}{\partial f_i} = \frac{\partial \gamma_i}{\partial f_i} A_i(P_i; f_i; w_i) + \gamma_i \frac{\partial A_i(P_i; f_i; w_i)}{\partial f_i} = 0, \tag{28}$$

and

$$\frac{\partial \pi_i}{\partial w_i} = \frac{\partial \gamma_i}{\partial w_i} A(P_i; f_i; w_i) + \gamma_i \frac{\partial A(P_i; f_i; w_i)}{\partial w_i} = 0. \tag{29}$$

I start by equation (27). From (2), I find that $\partial \gamma_i / \partial P_i = -1/2t$. As $\partial A(P_i; f_i; w_i)/\partial P_i = 1$, replacing in (27), I obtain

$$A(P_i; f_i; w_i) - 2t\gamma_i = 0. \tag{30}$$

Now consider equation (28). From (2), I find that $\partial \gamma_i / \partial f_i = -(1/2t)\partial C_i^*(\lambda_i^*)/\partial f_i$. Hence, using (24), I have that $\dfrac{\partial \gamma_i}{\partial f_i} = -\dfrac{1}{2t}\displaystyle\int_{\lambda_i^*}^{\bar{\lambda}} \frac{F(T)}{T} dT$. As

$$\frac{\partial A(P_i; f_i; w_i)}{\partial f_i} = \int_{\lambda_i^*}^{\bar{\lambda}} \frac{F(T)}{T} dT - \frac{\partial \lambda_i^*}{\partial f_i} \frac{F(\lambda_i^*)}{\lambda_i^*}(f_i + a^C - c_I) + \frac{\partial n_i^*}{\partial f_i}((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W) - k\varphi_i F(\lambda_i^*)\frac{\partial \lambda_i^*}{\partial f_i},$$

replacing in (28), and using (30), after a simplification by $\gamma_i > 0$, equation (28) can be rewritten as

$$\frac{\partial A(P_i; f_i; w_i)}{\partial f_i} - \frac{\partial C_i^*(\lambda_i^*)}{\partial f_i} = 0.$$

Hence, we have that

$$\frac{\partial \lambda_i^*}{\partial f_i} \frac{F(\lambda_i^*)}{\lambda_i^*}(f_i + a^C - c_I) - \frac{\partial n_i^*}{\partial f_i}\left[(1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W\right] + k\varphi_i F(\lambda_i^*)\frac{\partial \lambda_i^*}{\partial f_i} = 0. \tag{31}$$

From (17), I obtain that

$$\frac{\partial n_i^*}{\partial f_i} = \sqrt{\frac{rS(\lambda_i^*)}{2((1-\varphi_i)w_i + b)}} \times \frac{F(\lambda_i^*)}{2S(\lambda_i^*)} \frac{\partial \lambda_i^*}{\partial f_i}.$$

Replacing for this expression in (31), I get the equation that defines the card fee

$$-\frac{\partial \lambda_i^*}{\partial f_i} \frac{F(\lambda_i^*)}{\lambda_i^*}\left[f_i + a^C - c_I - \frac{n_i^* \lambda_i^*}{2S(\lambda_i^*)}((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W) + k\varphi_i \lambda_i^*\right] = 0.$$

Therefore,

$$f_i + a^C - c_I - \frac{n_i^*\lambda_i^*}{2S(\lambda_i^*)}((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W) + k\varphi_i\lambda_i^* = 0. \tag{32}$$

I now study equation (29). From (2), I obtain $\partial\gamma_i/\partial w_i = -(1/2t)\partial C_i^*(\lambda_i^*)/\partial w_i$. Using (23), I find that $\partial\gamma_i/\partial w_i = -(1-\varphi_i)n_i^*/(2t)$. As

$$\frac{\partial A(P_i;f_i;w_i)}{\partial w_i} = -\frac{\partial\lambda_i^*}{\partial w_i}\frac{F(\lambda_i^*)}{\lambda_i^*}(f_i+a^C-c_I)+\frac{\partial n_i^*}{\partial w_i}((1-\varphi_i)w_i-\varphi_i c_W-(1-\varphi_i)a^W)+(1-\varphi_i)n_i^*-k\varphi_i F(\lambda_i^*)\frac{\partial\lambda_i^*}{\partial w_i},$$

replacing in (29) and using (30), equation (29) can be rewritten as

$$\frac{\partial\lambda_i^*}{\partial w_i}\frac{F(\lambda_i^*)}{\lambda_i^*}(f_i + a^C - c_I) - \frac{\partial n_i^*}{\partial w_i}((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W) + k\varphi_i F(\lambda_i^*)\frac{\partial\lambda_i^*}{\partial w_i} = 0. \tag{33}$$

From (17), I have $\dfrac{\partial n_i^*}{\partial w_i} = \dfrac{n_i^* F(\lambda_i^*)}{2S(\lambda_i^*)}\dfrac{\partial\lambda_i^*}{\partial w_i} - \dfrac{(1-\varphi_i)n_i^*}{2((1-\varphi_i)w_i+b)}$. Replacing in (33), I obtain

$$-\frac{\partial\lambda_i^*}{\partial w_i}\frac{F(\lambda_i^*)}{\lambda_i^*}\left[f_i + a^C - c_I - \frac{n_i^*\lambda_i^*}{2S(\lambda_i^*)}((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W) + k\varphi_i\lambda_i^*\right] \tag{34}$$

$$= \frac{-n_i^*\lambda_i^*((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W)}{2((1-\varphi)w_i+b)}. \tag{35}$$

I denote by $(P^*;f^*;w^*)$ the candidate symmetric equilibrium solution of (27), (28), and (29). As $f + a^C - c_I + k\lambda_i^* = n_i^*\lambda_i^*((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W)/2S(\lambda_i^*)$, by equation (32) and from equation (34), the withdrawal fee is

$$w_i^* = a^W + \varphi_i c_W/(1-\varphi_i).$$

Hence, the card fee is

$$f_i^* = c_I - a^C - k\varphi_i\lambda_i^*.$$

From (30), the deposit fee then verifies

$$P_i^* = t + \left[2\left(M_{FC}^i\right)^* + \left(M_{FC}^j\right)^* - 2\left(M_{HC}^i\right)^* - \left(M_{HC}^j\right)^* + C_j^*(\lambda_j^*) - C_i^*(\lambda_i^*)\right]/3.$$

Appendix B-2: Second-order conditions. I provide here the conditions under which the second-order conditions are verified at $p^* = (P_i^*;f_i^*;w_i^*)$ by computing the coefficients of the Hessian matrix.

Denoting the Hessian matrix at $p^* = (P_i^*; f_i^*; w_i^*)$ by $H = \begin{pmatrix} a_1 & b & c \\ b & a_2 & d \\ c & d & a_3 \end{pmatrix}$, the second order conditions are verified if $a_1 \leq 0$, $a_2 \leq 0$, $a_1 a_2 - b^2 \geq 0$, $a_1 a_3 - c^2 \geq 0$, $a_3 a_2 - d^2 \geq 0$ and $\det H \leq 0$ (See hereafter). If these conditions are verified, this proves that the Hessian matrix is semi-definite negative at $p^* = (P_i^*; f_i^*; w_i^*)$.

I start by computing a_1 using the first equation (27), which defines the deposit fee. I have

$$\frac{\partial^2 \pi_i}{\partial^2 P_i} = \frac{\partial^2 \gamma_i}{\partial^2 P_i} A + 2 \frac{\partial \gamma_i}{\partial P_i} \frac{\partial A}{\partial P_i} + \gamma_i \frac{\partial^2 A}{\partial^2 P_i}.$$

As $\partial \gamma_i / \partial P_i = -1/2t$, $\partial^2 \gamma_i / \partial^2 P_i = 0$. As $\partial A / \partial P_i = 1$, $\partial^2 A / \partial^2 P_i = 0$. Hence,

$$a_1 = \left. \frac{\partial^2 \pi_i}{\partial^2 P_i} \right|_{p^*} = -\frac{1}{t} \leq 0.$$

I now compute the coefficient b. The derivative of the first equation with respect to f_i yields

$$\frac{\partial^2 \pi_i}{\partial P_i \partial f_i} = \frac{\partial^2 \gamma_i}{\partial P_i \partial f_i} A + 2 \frac{\partial \gamma_i}{\partial P_i} \frac{\partial A}{\partial f_i} + \gamma_i \frac{\partial^2 A}{\partial P_i \partial f_i}.$$

As $\partial \gamma_i / \partial P_i = -1/2t$ and $\partial A / \partial P_i = 1$, then $\partial^2 \gamma_i / \partial P_i \partial f_i = 0$ and $\partial^2 A / \partial P_i \partial f_i = 0$. At $p^* = (P_i^*; f_i^*; w_i^*)$, from (31), $\left. \frac{\partial A}{\partial f_i} \right|_{p^*} = \left. \frac{\partial C}{\partial f_i} \right|_{p^*}$. Hence,

$$\left. \frac{\partial^2 \pi_i}{\partial P_i \partial f_i} \right|_{p^*} = -\frac{1}{t} \left. \frac{\partial C}{\partial f_i} \right|_{p^*}.$$

From Lemma (2), I have that $\partial C_i / \partial f_i \geq 0$, hence,

$$b = \left. \frac{\partial^2 \pi_i}{\partial P_i \partial f_i} \right|_{p^*} \leq 0.$$

Similarly, I can prove that c is negative, that is,

$$c = \left. \frac{\partial^2 \pi_i}{\partial P_i \partial w_i} \right|_{p^*} = -\frac{1}{t} \left. \frac{\partial C}{\partial w_i} \right|_{p^*} \leq 0.$$

I now study the second equation (28), which defines the card fee, in order to compute a_2. I have

$$\frac{\partial^2 \pi_i}{\partial^2 f_i} = \frac{\partial^2 \gamma_i}{\partial^2 f_i} A + 2 \frac{\partial \gamma_i}{\partial f_i} \frac{\partial A}{\partial f_i} + \gamma_i \frac{\partial^2 A}{\partial^2 f_i}.$$

As $\frac{\partial \gamma_i}{\partial f_i} = \frac{-1}{2t}\frac{\partial C_i}{\partial f_i}$, $\frac{\partial^2 \gamma_i}{\partial^2 f_i} = -\frac{1}{2t}\frac{\partial^2 C_i}{\partial^2 f_i}$. From (24), I obtain that $\frac{\partial^2 \gamma_i}{\partial^2 f_i} = \frac{1}{2t}\frac{\partial \lambda_i^*}{\partial f_i}\frac{F(\lambda_i^*)}{\lambda_i^*}$. As

$$\frac{\partial A(P_i; f_i; w_i)}{\partial f_i} = \int_{\lambda_i^*}^{\overline{\lambda}} \frac{F(T)}{T}dT - \frac{\partial \lambda_i^*}{\partial f_i}\frac{F(\lambda_i^*)}{\lambda_i^*}(f_i + a^C - c_I) + \frac{\partial n_i^*}{\partial f_i}((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W) - \frac{\partial \lambda_i^*}{\partial f_i}F(\lambda_i^*)k\varphi_i,$$

I can compute the second derivative of A with respect to f_i at p^*. This yields[49]

$$\left.\frac{\partial^2 A(P_i; f_i; w_i)}{\partial^2 f_i}\right|_{p^*} = -\left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\frac{F(\lambda_i^*)}{\lambda_i^*}(2 + k\varphi_i \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}).$$

Hence,

$$\left.\frac{\partial^2 \pi_i}{\partial^2 f_i}\right|_{p^*} = \frac{-1}{t}\left(\left.\frac{\partial C_i}{\partial f_i}\right|_{p^*}\right)^2 - \gamma_i \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\frac{F(\lambda_i^*)}{\lambda_i^*}\left(1 + k\varphi_i \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right).$$

As $\partial \lambda_i^*/\partial f_i \geq 0$, we have that

$$a_2 = \left.\frac{\partial^2 \pi_i}{\partial^2 f_i}\right|_{p^*} \leq 0.$$

I now compute the coefficient d using the cross derivative of π_i with respect to w_i and f_i. This yields

$$\frac{\partial^2 \pi_i}{\partial f_i \partial w_i} = \frac{\partial^2 \gamma_i}{\partial f_i \partial w_i}A + \frac{\partial \gamma_i}{\partial f_i}\frac{\partial A}{\partial w_i} + \frac{\partial \gamma_i}{\partial w_i}\frac{\partial A}{\partial f_i} + \gamma_i \frac{\partial^2 A}{\partial f_i \partial w_i}.$$

The cross derivative of A with respect to f_i and w_i is

$$\left.\frac{\partial^2 A(P_i; f_i; w_i)}{\partial f_i \partial w_i}\right|_{p^*} = -\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\frac{F(\lambda_i^*)}{\lambda_i^*}(1 + k\varphi_i \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}) + (1-\varphi_i)\left.\frac{\partial n_i^*}{\partial f_i}\right|_{p^*}.$$

I also have $\frac{\partial^2 \gamma_i}{\partial f_i \partial w_i} = -\frac{1}{2t}\frac{\partial^2 C_i^*}{\partial w_i \partial f_i}$. As $\frac{\partial C_i^*}{\partial w_i} = (1-\varphi_i)n_i^*$, then $\frac{\partial^2 \gamma_i}{\partial f_i \partial w_i} = -\frac{1}{2t}(1-\varphi)\frac{\partial n_i^*}{\partial f_i}$. Hence, from (30), after some simplifications,

$$d = \left.\frac{\partial^2 \pi_i}{\partial f_i \partial w_i}\right|_{p^*} = \frac{-1}{t}\left(\left.\frac{\partial C_i}{\partial f_i}\right|_{p^*}\right)\left(\left.\frac{\partial C_i}{\partial w_i}\right|_{p^*}\right) - \gamma_i \left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\frac{F(\lambda_i^*)}{\lambda_i^*}(1 + k\varphi_i \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}).$$

From Lemma 1, this proves that $d \leq 0$.

Finally, I compute a_3, by studying the second derivative of π_i with respect to w_i. Using (29), I obtain,

$$\frac{\partial^2 \pi_i}{\partial^2 w_i} = \frac{\partial^2 \gamma_i}{\partial^2 w_i}A + 2\frac{\partial \gamma_i}{\partial w_i}\frac{\partial A}{\partial w_i} + \gamma_i \frac{\partial^2 A}{\partial^2 w_i}.$$

[49] I do not provide all the detail of the computation here. The reader can usefully notice that at p^*, $f_i + a^C - c_I = -\lambda_i^* k\varphi_i$ and that $(1-\varphi_i)w_i = \varphi_i c_W + (1-\varphi_i)a^W$.

As

$$\frac{\partial A(P_i;f_i;w_i)}{\partial w_i} = -\frac{\partial \lambda_i^*}{\partial w_i}\frac{F(\lambda_i^*)}{\lambda_i^*}(f_i+a^C-c_I)+\frac{\partial n_i^*}{\partial w_i}((1-\varphi_i)w_i-\varphi_i c_W-(1-\varphi_i)a^W)+(1-\varphi_i)n_i^*-\frac{\partial \lambda_i^*}{\partial w_i}F(\lambda_i^*)k\varphi_i,$$

I have

$$\left.\frac{\partial^2 A(P_i;f_i;w_i)}{\partial^2 w_i}\right|_{p^*} = \frac{F(\lambda_i^*)}{\lambda_i^*}k\varphi_i\left(\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\right)^2 + 2(1-\varphi_i)\left.\frac{\partial n_i^*}{\partial w_i}\right|_{p^*}.$$

As $\frac{\partial \gamma_i}{\partial w_i} = \frac{-1}{2t}\frac{\partial C_i}{\partial w_i}$, $\frac{\partial^2 \gamma_i}{\partial^2 w_i} = \frac{-1}{2t}\frac{\partial^2 C_i}{\partial^2 w_i}$. From (23), I obtain $\frac{\partial^2 \gamma_i}{\partial^2 w_i} = -\frac{1-\varphi_i}{2t}\frac{\partial n_i^*}{\partial w_i}$. Hence,

$$\left.\frac{\partial^2 \pi_i}{\partial^2 w_i}\right|_{p^*} = (1-\varphi_i)\gamma_i\left.\frac{\partial n_i^*}{\partial w_i}\right|_{p^*} - \frac{1}{t}\left(\left.\frac{\partial C_i}{\partial w_i}\right|_{p^*}\right)^2 - \gamma_i\frac{F(\lambda_i^*)}{\lambda_i^*}k\varphi_i\left(\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\right)^2.$$

As $\partial n_i^*/\partial w_i \leq 0$ and $\partial \lambda_i^*/\partial w_i \leq 0$, we have $a_3 \leq 0$.

I am now able to determine the conditions under which H is semi definitive negative. I have $a_1 = -1/t \leq 0$, and $a_2 = \partial^2 \pi_i/\partial^2 f_i \leq 0$. Hence, the first two conditions are verified. I also have

$$a_1a_2 - b^2 = \frac{\gamma_i F(\lambda_i^*)}{t\lambda_i^*}\left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\left(1+\varphi_i k\left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right) \geq 0,$$

$$a_1a_3 - c^2 = \frac{-(1-\varphi_i)\gamma_i}{t}\left(\left.\frac{\partial n_i^*}{\partial w_i}\right|_{p^*}\right) + \gamma_i\frac{F(\lambda_i^*)}{t\lambda_i^*}k\varphi_i\left(\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\right)^2 \geq 0,$$

$$\begin{aligned}
a_3a_2 - d^2 \;=\;& -\frac{1}{t}\left(\left.\frac{\partial C_i}{\partial f_i}\right|_{p^*}\right)^2(1-\varphi_i)\gamma_i\left.\frac{\partial n_i^*}{\partial w_i}\right|_{p^*} + \frac{\gamma_i}{t}\frac{F(\lambda_i^*)}{\lambda_i^*}k\varphi_i\left(\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\right)^2\left(\left.\frac{\partial C_i}{\partial f_i}\right|_{p^*}\right)^2 \\
& -(\gamma_i)^2\left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\frac{F(\lambda_i^*)}{\lambda_i^*}\left(1+\varphi_i k\left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right)(1-\varphi_i)\left.\frac{\partial n_i^*}{\partial w_i}\right|_{p^*} \\
& +\frac{\gamma_i}{t}\left(\left.\frac{\partial C_i}{\partial w_i}\right|_{p^*}\right)^2\left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\frac{F(\lambda_i^*)}{\lambda_i^*}\left(1+\varphi_i k\left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right) \\
& -\frac{2\gamma_i}{t}\left.\frac{\partial C_i}{\partial f_i}\right|_{p^*}\left.\frac{\partial C_i}{\partial w_i}\right|_{p^*}\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\frac{F(\lambda_i^*)}{\lambda_i^*}\left(1+\varphi_i k\left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right) \\
& -(\gamma_i)^2\left(\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\right)^2\left(\frac{F(\lambda_i^*)}{\lambda_i^*}\right)^2\left(1+\varphi_i k\left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right).
\end{aligned}$$

All the terms except the last one are positive. We have to assume that the first terms are high enough, such that $a_3a_2 - d^2 \geq 0$. Finally,

$$\det H = \frac{(\gamma_i)^2 \, F(\lambda_i^*)}{t\lambda_i^*} \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*} \left(1 + \varphi_i k \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right) \left[(1-\varphi_i) \left.\frac{\partial n_i^*}{\partial w_i}\right|_{p^*} \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*} + \frac{F(\lambda_i^*)}{\lambda_i^*} \left(\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\right)^2\right].$$

We have $\det H \leq 0$ if and only if

$$(1-\varphi_i) \left.\frac{\partial n_i^*}{\partial w_i}\right|_{p^*} \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*} + \frac{F(\lambda_i^*)}{\lambda_i^*} \left(\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\right)^2 \leq 0.$$

I will assume that the distribution of transaction prices is chosen such that this condition is verified. For instance, it is possible to prove that, with the function F of Appendix A-2, this condition is verified. Hence, the second order conditions of profit maximisation are satisfied in this special case.

Appendix B-3: An example with the function $F(T) = 2ST$ with $T \in [0;1]$. From Appendix A-2, if $f_i > \sqrt{r((1-\varphi_i)w_i + b)/2S}$, we have

$$n_i^* = \lambda_i^* \sqrt{\frac{rS}{2((1-\varphi_i)w_i + b)}},$$

and

$$\lambda_i^* = \frac{1}{v_i - \beta(a^C + c_A)} \left(f_i - \sqrt{\frac{r((1-\varphi_i)w_i + b)}{2S}}\right).$$

Replacing for the transaction fees obtained in Proposition 2, it can be easily shown that

$$\begin{aligned} f_i^* &= \frac{v_i - \beta(a^C + c_A)}{v_i - \beta(a^C + c_A) + k\varphi_i} \left(c_I - a^C + \sqrt{\frac{r(\varphi_i c_W + (1-\varphi_i)a^W + b)}{2S}}\right), \\ \lambda_i^* &= \frac{1}{v_i - \beta(a^C + c_A) + k\varphi_i} \left(c_I - a^C - k\varphi_i \sqrt{\frac{r(\varphi_i c_W + (1-\varphi_i)a^W + b)}{2S}}\right), \end{aligned}$$

and

$$\widehat{a_i^C} = c_I - \sqrt{\frac{r(\varphi_i c_W + (1-\varphi_i)a^W + b)}{2S}}.$$

The expression of banks' margins at the equilibrium of stage 2 is

$$(M_{HC}^i)^* = -2Sk\varphi_i(\lambda_i^*)^2,$$

and

$$(M_{FC}^i)^* = n_j^*(1-\varphi_j)(a^W - c_W) - kS(1-\varphi_j)(\lambda_j^*)^2.$$

Appendix C: Proof of Lemma 4 and 5.

Appendix C-1: proof of Lemma 4. I check ex post the conditions under which consumers use both cash and cards.

First, let us determine the conditions under which consumers do not use cash at the equilibrium. From Proposition 1, if $f_i^* \leq l\sqrt{r((1-\varphi_i)w_i^* + b)/2}$, the consumer pays by card all his transactions. I now prove that f_i^* is decreasing with a^C. As $f_i^* = c_I - a^C - \varphi_i k \lambda_i^*$, we have that

$$\frac{\partial f_i^*}{\partial a^C} = -1 - k\varphi_i \left[\frac{\partial \lambda_i^*}{\partial f_i} \frac{\partial f_i^*}{\partial a^C} + \frac{\partial \lambda_i^*}{\partial m} \right].$$

Hence,

$$\frac{\partial f_i^*}{\partial a^C} = \frac{-1 - k\varphi_i(\partial \lambda_i^*/\partial m)}{1 + k\varphi_i(\partial \lambda_i^*/\partial f_i)} \tag{36}$$

As $\partial \lambda_i^*/\partial f_i \geq 0$ and $\partial \lambda_i^*/\partial m \geq 0$ from Lemma 1 if β is sufficiently small, we conclude that the card fee decreases with the interchange fee on card payment if the surcharge rate is sufficiently small, that is $\partial\ f_i^*/\partial a^C \leq 0$. Hence, there exists a level of interchange fee $\widehat{a_i^C}(a^W)$ such that $f_i^* > l\sqrt{r((1-\varphi_i)w_i^* + b)/2}$ if $a^C < \widehat{a_i^C}(a^W)$ and $f_i^* \leq l\sqrt{r((1-\varphi_i)w_i^* + b)/2}$ otherwise. As a consequence, using the result of Proposition 1, we conclude that the consumers of bank i do not use cash if $a^C \geq \widehat{a_i^C}(a^W)$.

Second, let us determine the conditions under which consumers do not use cards at the equilibrium. Assume that $a^W \geq c_W$. Then $w_i^* \geq c_W$. Because of assumption (A4), we have

$$c_I < \overline{\lambda}(v_i + \sqrt{r((1-\varphi_i)w_i^* + b)/2S(\overline{\lambda})}).$$

Hence, it follows that $f_i^* + a^C + k\varphi_i\lambda_i^* < \overline{\lambda}(v_i - \beta(a^C + c_A) + \sqrt{r((1-\varphi_i)w_i^* + b)/2S(\overline{\lambda})})$ if β is sufficiently small. As $a^C \geq 0$ and $k\varphi_i\lambda_i^* \geq 0$, we have that $f_i^* < \overline{\lambda}(v_i + \sqrt{r((1-\varphi_i)w_i^* + b)/2S(\overline{\lambda})})$. From Proposition 1, I can conclude that it is never optimal for the consumers to use only cash if banks price the transactions at the average cost.

Hence, if β is sufficiently small, consumers use a combination of cash and card payments at the equilibrium if and only if $a^C < \widehat{a_i^C}(a^W)$.

Appendix C-2: proof of Lemma 5. From Lemma 1, the threshold λ_i^* above which consumers pay by card, and the number of withdrawals, n_i^*, increase with the card fee and decrease with the withdrawal fee. From Lemma 4, the card fee decreases with the interchange fee if β is sufficiently small.

The threshold above which the consumers pay by card λ_i^* is indirectly related to the interchange fee on card payments through the merchant fee and the card fee, that is

$$\frac{d\lambda_i^*}{da^C} = \frac{\partial\lambda_i^*}{\partial m}\frac{\partial m}{\partial a^C} + \frac{\partial\lambda_i}{\partial f_i}\frac{\partial f_i^*}{\partial a^C}.$$

As $\partial\lambda_i^*/\partial m \geq 0$, $m = a^C + c_A$, $\partial\lambda_i^*/\partial f_i \geq 0$ and $\partial f_i/\partial a^C \leq 0$ if β is sufficiently small, the threshold value decreases with the interchange fee on card payments, that is $\partial\lambda_i^*/\partial a^C \leq 0$. As $\partial n_i/\partial\lambda_i \geq 0$, we conclude that the number of withdrawals also decreases with the interchange fee on card payments, that is $\partial n_i^*/\partial a^C \leq 0$.

As $w_i^* = (\varphi_i/(1-\varphi_i))c_W + a^W$, we have that $\partial w_i^*/\partial a^W \geq 0$. As the number of withdrawals decreases with the withdrawal fee, we conclude that the number of withdrawals decreases with the interchange fee on card payments, that is $\partial n_i^*/\partial a^W \leq 0$. Similarly, we have that $\partial\lambda_i^*/\partial a^W \leq 0$.

To sum up, the threshold λ_i^* and the number of withdrawals n_i^* decrease with the interchange fee on card payments and increase with the interchange fee on withdrawals.

I now prove that the transaction costs increase with the interchange fee on withdrawals, while they decrease with the interchange fee on card payments if β is sufficiently small. Using the envelop's theorem, I obtain the total derivative of the consumer's transaction costs with respect to the interchange fees, that is

$$\frac{dC_i^*}{da^C} = \frac{\partial C_i^*}{\partial f_i}\frac{\partial f_i^*}{\partial a^C} + \frac{\partial C_i^*}{\partial m}\frac{\partial m}{\partial a^C},$$

and

$$\frac{dC_i^*}{da^W} = \frac{\partial C_i^*}{\partial w_i}\frac{\partial w_i}{\partial a^W}.$$

As $m = a^C + c_A$, and $w_i = (\varphi_i/(1-\varphi_i))c_W + a^W$, and using (24), (26), and (23), the total derivatives of the costs with respect to the interchange fees are

$$\frac{dC_i^*}{da^C} = \frac{\partial f_i^*}{\partial a^C}\int\limits_{\lambda_i(a^C,a^W)}^{\bar{\lambda}} \frac{F(T)}{T}dT + \beta\int\limits_{\lambda_i(a^C,a^W)}^{\bar{\lambda}} F(T)dT, \tag{37}$$

and

$$\frac{dC_i^*}{da^W} = (1-\varphi_i)n_i^*. \tag{38}$$

If β is small, as $\partial f_i^*/\partial a^C \leq 0$, we have that $\partial C_i^*/\partial a^C \leq 0$ and $\partial C_i^*/\partial a^W \geq 0$. It follows that the transaction costs increase with the interchange fee on withdrawals, while they decrease with the interchange fee on card payments if β is sufficiently small.

Appendix D: user surplus and welfare.

Appendix D-1: impact of interchange fees on merchant surplus. Assume that consumers use both cash and cards to pay for their expenses. From (10), we have that

$$\frac{dMS}{da^C} = -(1-\beta)\int_{\lambda^*}^{\overline{\lambda}} F(T)dT - (c_M - (1-\beta)(a^C + c_A))\frac{\partial \lambda^*}{\partial a^C}F(\lambda^*),$$

and

$$\frac{dMS}{da^W} = -(c_M - (1-\beta)(a^C + c_A))\frac{\partial \lambda^*}{\partial a^W}F(\lambda^*).$$

As $\partial\lambda^*/\partial a^W \leq 0$, the merchant surplus is increasing with the interchange fee on withdrawals. A rise in the interchange fee on card payments has a positive effect on the volume of transaction that is paid by card, which impacts positively the merchant surplus, but a negative effect on the merchants' margin per transaction, which impacts negatively the merchant surplus.

If the consumers use only cards to pay for their expenses, the merchant surplus is

$$MS = S - (1-\beta)(a^C + c_A)S.$$

The merchant surplus decreases with the interchange fee on card payments. As there is no need to encourage consumers to substitute card payments for cash, the interchange fee on withdrawals has no impact on merchant surplus.

Appendix D-2: impact of interchange fees on consumer surplus. Assume that consumers use both cash and cards to pay for their expenses. We determine the derivatives of the deposit fee and of the consumer transaction costs with respect to the interchange fee on card payments. From (12), the derivatives with respect to a^C are

$$\frac{dP^*}{da^C} = \frac{\partial n^*}{\partial a^C}(1-\varphi)(a^W - c_W) - k(1-\varphi)\frac{\partial \lambda^*}{\partial a^C}F(\lambda^*) + k\varphi\frac{\partial \lambda^*}{\partial a^C}\int_{\lambda^*}^{\overline{\lambda}} \frac{F(T)}{T}dT,$$

and from (37),

$$\frac{dC^*}{da^C} = \frac{\partial f_i^*}{\partial a^C}\int_{\lambda^*}^{\overline{\lambda}} \frac{F(T)}{T}dT + \beta\int_{\lambda^*}^{\overline{\lambda}} F(T)dT.$$

We also determine the derivative of the deposit fee and of the consumer transaction costs with respect to the interchange fee on withdrawals. We obtain

$$\frac{dP^*}{da^W}=\frac{\partial n^*}{\partial a^W}(1-\varphi)(a^W-c_W)+(1-\varphi)n^*-k(1-\varphi)\frac{\partial \lambda^*}{\partial a^W}F(\lambda^*)+k\varphi\frac{\partial \lambda^*}{\partial a^W}\int_{\lambda^*}^{\overline{\lambda}}\frac{F(T)}{T}dT,$$

and from (38),

$$\frac{dC^*}{da^W}=(1-\varphi)n^*.$$

As $CS=V+B-S-(P^*+C^*)$, the derivatives of CS with respect to a^C and a^W are

$$\begin{aligned}\frac{dCS}{da^C} &= -\frac{\partial n^*}{\partial a^C}(1-\varphi)(a^W-c_W)+k(1-\varphi)\frac{\partial \lambda^*}{\partial a^C}F(\lambda^*)-k\varphi\frac{\partial \lambda^*}{\partial a^C}\int_{\lambda^*}^{\overline{\lambda}}\frac{F(T)}{T}dT\\ &\quad -\frac{\partial f_i^*}{\partial a^C}\int_{\lambda^*}^{\overline{\lambda}}\frac{F(T)}{T}dT-\beta\int_{\lambda^*}^{\overline{\lambda}}F(T)dT,\end{aligned}$$

and

$$\frac{dCS}{da^W}=-\frac{\partial n^*}{\partial a^W}(1-\varphi)(a^W-c_W)-2(1-\varphi)n^*+k(1-\varphi)\frac{\partial \lambda^*}{\partial a^W}F(\lambda^*)-k\varphi\frac{\partial \lambda^*}{\partial a^W}\int_{\lambda^*}^{\overline{\lambda}}\frac{F(T)}{T}dT.$$

Now suppose that consumers use only payment cards to pay for their expenses. Then we have

$$CS=V+B-S-(c_I-a^C)\int_{0}^{\overline{\lambda}}\frac{F(T)}{T}dT+vS-t-\beta\int_{\lambda^*}^{\overline{\lambda}}F(T)dT.$$

The consumer surplus increases with the interchange fee on card payments, as an increase in a^C lowers the cost of paying by card. The interchange fee on withdrawals has no impact on consumer surplus.

Appendix D-3: Interchange fees and total user surplus. Assume that there is an interior solution to the problem of profit maximisation, that we denote by $IF=\left((a^C)^P,(a^W)^P\right)$, such that consumers use a combination of cash and card payments to pay for their expenses. From (9) and (8), we have that

$$\left.\frac{dCS}{da^C}\right|_{IF}=-k\varphi\left.\frac{\partial \lambda^*}{\partial a^C}\right|_{IF}\int_{\lambda(IF)}^{\overline{\lambda}}\frac{F(T)}{T}dT-\frac{\partial f^*}{\partial a^C}\int_{\lambda(IF)}^{\overline{\lambda}}\frac{F(T)}{T}dT-\beta\int_{\lambda(IF)}^{\overline{\lambda}}F(T)dT,$$

and

$$\left.\frac{dCS}{da^W}\right|_{IF} = -(1-\varphi)n^* - k\varphi \left.\frac{\partial\lambda^*}{\partial a^W}\right|_{IF} \int\limits_{\lambda(IF)}^{\overline{\lambda}} \frac{F(T)}{T} dT.$$

As $\left.\frac{\partial\lambda^*}{\partial a^C}\right|_{IF} \leq 0$ and $\left.\frac{d\lambda^*}{df}\right|_{IF} \geq 0$, we have that $\left.\frac{dCS}{da^C}\right|_{IF} \geq 0$ if β is small enough. It follows that the profit maximising interchange fee on card payments is too low to maximise consumer surplus. If k is small and $\varphi \neq 1$, we also have that $\left.\frac{dCS}{da^W}\right|_{IF} \leq 0$. So the profit maximising interchange fee on withdrawals is too high to maximise consumer surplus.

As $\frac{dMS}{da^W} \geq 0$, the interchange fee on withdrawals is too low to maximise merchant surplus, as long as consumers still use cash to pay for their expenses. From Appendix D-1, we have that

$$\left.\frac{dMS}{da^C}\right|_{IF} = -(1-\beta) \int\limits_{\lambda(IF)}^{\overline{\lambda}} F(T)dT - (c_M - (1-\beta)((a^C)^P + c_A)) \left.\frac{\partial\lambda^*}{\partial a^C}\right|_{IF} F(\lambda(IF)).$$

If the value of the expenses paid by card is very high, then the profit maximising interchange fee on card payments is too high to maximise merchant surplus as $\left.\frac{dMS}{da^C}\right|_{IF} \leq 0$. If the value of card payments is very low, or if the cost of cash for merchants is very high, the profit maximising interchange fee on card payments is too low to maximise merchant surplus, as $\left.\frac{dMS}{da^C}\right|_{IF} \geq 0$.

Let us analyse determine how the profit maximising interchange fees compare with the interchange fees that maximise total user surplus. We have that

$$\begin{aligned} \left.\frac{dTUS}{da^C}\right|_{IF} &= -k\varphi \left.\frac{\partial\lambda^*}{\partial a^C}\right|_{IF} \int\limits_{\lambda(IF)}^{\overline{\lambda}} \frac{F(T)}{T} dT - \frac{\partial f^*}{\partial a^C} \int\limits_{\lambda(IF)}^{\overline{\lambda}} \frac{F(T)}{T} dT \\ &\quad - \int\limits_{\lambda(IF)}^{\overline{\lambda}} F(T)dT - (c_M - (1-\beta)((a^C)^P + c_A)) \left.\frac{\partial\lambda^*}{\partial a^C}\right|_{IF} F(\lambda^*). \end{aligned}$$

All the terms of the previous equation are positive, except $-\int\limits_{\lambda(IF)}^{\overline{\lambda}} F(T)dT$, which represents the value of card payments. If the cost of cash is low for banks and for merchants, and if the number of transactions paid by card (equal to $\int\limits_{\lambda(IF)}^{\overline{\lambda}} \frac{F(T)}{T} dT$) is low compared to the value of card payments (equal to $\int\limits_{\lambda(IF)}^{\overline{\lambda}} F(T)dT$), the interchange fee on card payments is too high to maximise total user surplus, as $\left.\dfrac{dTUS}{da^C}\right|_{IF} \leq 0$. The only case in which the profit maximising interchange fee on card payments may be too low to maximise total user surplus is the case in which the cost of cash is high for banks or merchants, and the value of card payments is low. We now study the case of the interchange fee on withdrawals. We have that

$$\left.\frac{dTUS}{da^W}\right|_{IF} = -(1-\varphi)n^* - k\varphi \left.\frac{\partial\lambda^*}{\partial a^W}\right|_{IF} \int_{\lambda(IF)}^{\overline{\lambda}} \frac{F(T)}{T} dT - (c_M - (1-\beta)((a^C)^P + c_A)) \left.\frac{\partial\lambda^*}{\partial a^W}\right|_{IF} F(\lambda^*).$$

All the terms of the previous equation are positive, except $-(1-\varphi)n^*$, which represents the volume of foreign withdrawals. If the benefit of accepting cards is small for merchants, that is if $(c_M - (1-\beta)((a^C)^P + c_A))$ is small, if the cost of cash for banks is not too high, and if the volume of foreign withdrawals is significant, the interchange fee on withdrawals is too high to maximise total user surplus. If k is high, and if the cost of cash for merchants is high, the interchange fee on withdrawals is too low to maximise the total user surplus.

Appendix D-4: Interchange fees and social welfare. We determine the derivatives of social welfare with respect to the interchange fees on card payments and cash withdrawals. We have

$$\frac{\partial W}{\partial a^W} = \underbrace{(1-\varphi)(a^W - c_W)\frac{\partial n^*}{\partial a^W}}_{A} - \underbrace{k\frac{\partial\lambda^*}{\partial a^W}\left((1-\varphi)F(\lambda^*) + \varphi\int_{\lambda^*}^{\overline{\lambda}}\frac{F(T)}{T}dT\right)}_{B} - \underbrace{(c_M - (1-\beta)(a^C + c_A))\frac{\partial\lambda^*}{\partial a^W}F(\lambda^*)}_{C} \tag{39}$$

and

$$\frac{\partial W}{\partial a^C} = \underbrace{(1-\varphi)(a^W - c_W)\frac{\partial n^*}{\partial a^C}}_{D} - \underbrace{k\frac{\partial\lambda^*}{\partial a^C}\left((1-\varphi)F(\lambda^*) + \varphi\int_{\lambda^*}^{\overline{\lambda}}\frac{F(T)}{T}dT\right)}_{E} + \underbrace{\frac{\partial f_i^*}{\partial a^C}\int_{\lambda^*}^{\overline{\lambda}}\frac{F(T)}{T}}_{F} \tag{40}$$

$$\underbrace{-\int_{\lambda^*}^{\overline{\lambda}} F(T)dT}_{G} - \underbrace{(c_M - (1-\beta)(a^C + c_A))\frac{\partial\lambda^*}{\partial a^C}F(\lambda^*)}_{H}. \tag{41}$$

Appendix D-5: ex-post conditions on Z such that it is always in the interest of the payment association that merchants accept cards. If the merchants refuse cards, banks' joint profit is

$$\pi_{nc} = t + 2n^*(1-\varphi)(a^W - c_W) - 2k(1-\varphi)S - Z.$$

Let us denote by $(a^W)_{nc}$ the joint profit maximising interchange fee on withdrawals when the merchants refuse cards. It can be easily checked that the joint profit maximising interchange fee on withdrawals is equal to the monopoly price. We also denote by $(a^W)_c$ and $(a^C)_c$ the joint profit maximising interchange fees if the merchants accept cards. It is always in the interest of the payment platform that merchants accept cards if

$$\pi_{nc}((a^W)_{nc}) \leq t + 2(n^*)_c(1-\varphi)((a^W)_c - c_W) - 2k(1-\varphi)S((\lambda^*)_c),$$

that is if (assuming that $\varphi < 1$)

$$\frac{Z}{2(1-\varphi)} \geq (n^*)_{nc}((a^W)_{nc} - c_W) - (n^*)_c((a^W)_c - c_W) - k(S - S((\lambda^*)_c).$$

If $(n^*)_c((a^W)_c - c_W) \geq (n^*)_{nc}((a^W)_{nc} - c_W)$, this condition is always true. Otherwise, either the costs of cash must be sufficiently high, or the desutility Z that is born by the payment association when the merchants refuse cards must be high.

Appendix E: Examples. In this Appendix, we give a few examples of market structures in several European countries. In the first column, I give the name of the entity that manages the ATM network. In the second column, I give the name of the largest payment card systems (in terms of transaction volume) that operate in the country. In the last column, I precise whether the payment card systems (PCS) choose multilateral interchange fees for card payments, and whether there are also multilateral or bilateral interchange fees on withdrawals. The letters AV mean that the interchange fee is an Ad Valorem tariff.

Country	ATM networks	PC Systems	Interchange fees?
Denmark	Sumclearing/PBS.	PBS.	ATMs: entry fee.
			PCS: No.
France	System "CB"	System "CB"	ATMs: Yes. PCS: Yes.
			- Bilateral component.
UK	Largest: Link, managed by "Vocalink".	Visa, MasterCard.	Link: Yes. PCS: Yes.
Germany	The "Cash pools"	Ec-Karte. POZ.	PCS: No.
Finland	Managed by "Automatia".	Pankkikortti System.	ATMs: no IF. Entry fee.
	(Owned by the 5 largest banks)		PCS: No IF.
Sweden	ATMs are installed and owned by banks.	Visa	ATMs: bilateral IF.
			PCS: Yes.
Norway	Managed by BankAxept	BankAxept	ATMs: entry fee+ MIF.
Portugal	Multibanco (managed by SIBS)	SIBS	PCS: Yes (AV)
Italy	Bancomat (managed by SIA)	Bancomat (SIA)	PCS: Yes (AV)
Belgium	ATMs managed by the banks.	Banksys	ATMs: bilateral IF.
	(Formerly owned by Banksys).		PCS: Yes.
Spain	ServiRed	ServiRed	PCS: Yes
	Red Euro 6000	Red Euro 6000	ATMs: Yes.
	Telebanco 4B	Telebanco 4B	
Netherlands	Agreement between Postbank & Equens	Equens/Interpay.	PCS: Bilateral IF.

Sources of the table: PSE Consulting, Groupement des Cartes Bancaires, Interim Report on Payment Cards (European Commission).

Appendix F: Extensions

Appendix F-1: Heterogeneous merchants. In this Appendix, I assume that there are several merchants from different industries who differ on their cost of accepting cash c_M, which is distributed on $[0, 1]$ according to a uniform distribution. The share α of merchants who accept cards is given by

$$\alpha = P(c_M \geq (1-\beta)(a^C + c_A)) = 1 - (1-\beta)(a^C + c_A).$$

When the consumer decides how often to withdraw from the ATM network, he takes into account the fact that the debit card as a probability α to be accepted as a means of payment.

Hence, the consumer's transaction costs are

$$C_i(\lambda_i, n_i) = \frac{r}{2n_i}\left(\alpha S(\lambda_i) + (1-\alpha)\, S\right) + n_i((1-\varphi_i)w_i + b) + \alpha f_i \int_{\lambda_i}^{\overline{\lambda}} \frac{F(T)}{T} dT + \alpha(\beta m - v_i)(S - S(\lambda_i)).$$

The consumer minimizes his transaction costs as in the main model, which gives the equations that define the threshold λ and the number of withdrawals:

$$n_i^* = \sqrt{\frac{r\left(\alpha S(\lambda_i^*) + (1-\alpha)\, S\right)}{2((1-\varphi_i)w_i + b)}},$$

and

$$\lambda_i^* \sqrt{\frac{r((1-\varphi_i)w_i + b)}{2\left(\alpha S(\lambda_i^*) + (1-\alpha)\, S\right)}} - f_i + (v_i - \beta m)\lambda_i^* = 0.$$

Using the same method as in the main model, it is possible to prove that n_i^* and λ_i^* decrease with the share of merchants who accept cards α.

At stage 2, the issuing banks choose the deposit fee and the transaction fees that maximise their profit. The margins made on home and foreign consumers are modified as follows:

$$M_{HC}^i = (f_i + a^C - c_I)\alpha \int_{\lambda_i^*}^{\overline{\lambda}} \frac{F(T)}{T} dT + n_i^*((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W) - k\varphi_i \left[\alpha S(\lambda_i^*) + (1-\alpha)S\right],$$

and

$$n_j^*(1-\varphi_j)(a^W - c_W) - k(1-\varphi_j)\left[\alpha S(\lambda_j^*) + (1-\alpha)S\right].$$

The profit maximising transaction fees are the same as in the main model, that is

$$f_i = c_I - a^C - k\lambda_i^*,$$

and

$$w_i = \varphi_i c_W + (1-\varphi_i)a^W.$$

The deposit fee now depends on the share of merchants who accept cards, that is

$$P_i = t + n^*(1-\varphi)(a^W - c_W) - k(1-2\varphi)\left(\alpha S(\lambda^*) + (1-\alpha)\, S\right) + k\varphi\lambda^* \int_{\lambda^*}^{\overline{\lambda}} \frac{F(T)}{T} dT.$$

If the consumers tend to withdraw more from their home bank ($1 - 2\varphi \leq 0$), the deposit fee increases when the share of merchants who accept cards increase. The joint profit in a symmetric equilibrium is given by

$$\pi = t + 2n^*(1-\varphi)(a^W - c_W) - 2k(1-\varphi)\left[\alpha S(\lambda^*) + (1-\alpha)S\right].$$

When they choose the profit maximising interchange fee on card payments, the issuers now take into account the fact that a higher interchange fee on card payments increases the probability that the debit card is refused by the merchants, and thereby increases the cost of cash. If there is an interior solution, the first order conditions of joint profit maximisation become:

$$\frac{\partial \pi}{\partial a^W} = 2n^*(1-\varphi) + 2\frac{\partial n^*}{\partial a^W}(1-\varphi)(a^W - c_W) - 2k\alpha(1-\varphi)\frac{\partial \lambda^*}{\partial a^W}F(\lambda^*) = 0,$$

and

$$\frac{\partial \pi}{\partial a^W} = 2\frac{\partial n^*}{\partial a^C}(1-\varphi)(a^W - c_W) - 2k\alpha(1-\varphi)\frac{\partial \lambda^*}{\partial a^C}F(\lambda^*) + 2k(1-\varphi)\frac{\partial \alpha^*}{\partial a^C}\left(S - S(\lambda^*)\right) = 0.$$

Hence, if there is an interior solution, the joint profit maximising interchange fees verify

$$\frac{a^W - c_W}{a^W} = \frac{1}{\epsilon} + \frac{k\alpha^* F(\lambda^*)}{a^W}\frac{\partial \lambda^*/\partial a^W}{\partial n^*/\partial a^W},$$

and

$$\frac{\partial n^*}{\partial a^C}(a^W - c_W) = k\alpha\frac{\partial \lambda^*}{\partial a^C}F(\lambda^*) - k\frac{\partial \alpha^*}{\partial a^C}(S - S(\lambda^*)).$$

The mark-up that is charged above the marginal cost for the interchange fees on withdrawals now depends on the percentage of merchants who accept cards. Banks decide to impose a higher mark-up on withdrawals when card acceptance is wide, so as to encourage consumers to withdraw less cash, in order to reduce the variable costs of cash. However, issuing banks take also into account the fact that higher interchange fees on card payments may also increase the costs of cash, as they may reduce card acceptance.

We now analyse the impact of interchange fees on total user surplus. It is complex to analyse whether higher merchant acceptance of payment cards increases or reduces the consumers' costs, as consumer adjust their payments decision to internalize the share of merchants who accept cards. Higher interchange fees on card payments do not necessarily increase consumer surplus if the card fee becomes lower, but the share of merchants who accept cards is reduced. The derivatives of the consumer surplus at the profit maximising interchange fees are respectively

$$\left.\frac{dCS}{da^C}\right|_{IF} = -k\alpha\varphi \left.\frac{\partial\lambda^*}{\partial a^C}\right|_{IF} \int\limits_{\lambda(IF)}^{\overline{\lambda}} \frac{F(T)}{T}dT - \alpha\frac{\partial f^*}{\partial a^C} \int\limits_{\lambda(IF)}^{\overline{\lambda}} \frac{F(T)}{T}dT$$
$$-\beta\alpha \int\limits_{\lambda(IF)}^{\overline{\lambda}} F(T)dT - \left.\frac{\partial C}{\partial \alpha}\right|_{IF} \left.\frac{\partial \alpha}{\partial a^C}\right|_{IF},$$

and

$$\left.\frac{dCS}{da^W}\right|_{IF} = -(1-\varphi)n^* - k\alpha\varphi \left.\frac{\partial\lambda^*}{\partial a^W}\right|_{IF} \int\limits_{\lambda(IF)}^{\overline{\lambda}} \frac{F(T)}{T}dT.$$

The difference between this expression and the expression obtained under merchants' homogeneity is due to the presence of α in the previous equations. The consumer surplus decreases with the interchange fee on withdrawals if the cost of cash is low or if merchant acceptance of payment cards is very limited. If the consumers' transaction costs do not increase too much when card acceptance is reduced, the profit maximising interchange fee on card payments is too low to maximise consumer surplus as under merchants' homogeneity.

To evaluate the derivative of the merchant surplus at the profit maximising interchange fee, we take into account the fact that merchants are now separated into two categories: the share of merchants who accept cards, and cash only merchants. The surplus of the merchants who accept cards is not affected by the impacted by the variations of interchange fees. The effect of a rise in the interchange fees is exactly the same as in section 3 for the merchants who accept cards. The difference is that a higher interchange fee on card payment reduces the share of merchants who accept cards. The derivative of the total user surplus at the profit maximising interchange fee is now expressed as follows

$$\left.\frac{dTUS}{da^C}\right|_{IF} = -k\alpha\varphi \left.\frac{\partial\lambda^*}{\partial a^C}\right|_{IF} \int\limits_{\lambda(IF)}^{\overline{\lambda}} \frac{F(T)}{T}dT - \alpha\frac{\partial f^*}{\partial a^C} \int\limits_{\lambda(IF)}^{\overline{\lambda}} \frac{F(T)}{T}dT - \left.\frac{\partial C}{\partial \alpha}\right|_{IF} \left.\frac{\partial \alpha}{\partial a^C}\right|_{IF}$$
$$-\alpha \int\limits_{\lambda(IF)}^{\overline{\lambda}} F(T)dT - \alpha(c_M - (1-\beta)((a^C)^P + c_A)) \left.\frac{\partial\lambda^*}{\partial a^C}\right|_{IF} F(\lambda^*)$$
$$-\left.\frac{\partial \alpha}{\partial a^C}\right|_{IF} \left[c_M - (1-\beta)(a^C + c_A)\right](S - S(\lambda^*)),$$

and

$$\left.\frac{dTUS}{da^W}\right|_{IF} = -(1-\varphi)n^* - k\alpha\varphi \left.\frac{\partial\lambda^*}{\partial a^W}\right|_{IF} \int\limits_{\lambda(IF)}^{\overline{\lambda}} \frac{F(T)}{T} dT - \alpha(c_M - (1-\beta)((a^C)^P + c_A)) \left.\frac{\partial\lambda^*}{\partial a^W}\right|_{IF} F(\lambda^*).$$

Merchants' heterogeneity modifies the results of Proposition 4 as follows. Notice that $\left.\frac{dTUS}{da^C}\right|_{IF}$ is positive if merchants' acceptance of payment cards (the parameter α) is very low, as $\left.\frac{\partial\alpha}{\partial a^C}\right|_{IF}$ is negative. Hence, as we assumed that TUS is concave, the interchange fee on card payments may be too low to maximise total user surplus if merchants' acceptance is low. If merchants' acceptance of payment cards is high, and if a lower interchange fee on card payments does not impact merchants' acceptance too much, the interchange fee on card payments is too high to maximise total user surplus. The only difference between $\left.\frac{dTUS}{da^W}\right|_{IF}$ and the expression obtained in Appendix D-3 is related to the presence of the parameter α. Hence, Proposition 4 could be restated as follows: " the profit maximising interchange fee on withdrawals is too high to maximise the total user surplus if the costs of cash are small for banks, if merchants' acceptance of payment cards is low, and if the volume of foreign withdrawals is relatively high.

Appendix F-2: Asymmetries between issuers. In this case, we have that $\left(M_{FC}^2\right)^* = \left(M_{HC}^2\right)^* = 0$. Hence, the deposit fees are

$$P_1^* = t + \left[2\left(M_{FC}^1\right)^* - 2\left(M_{HC}^1\right)^* + C_2^*(\lambda_2^*) - C_1^*(\lambda_1^*)\right]/3,$$

and

$$P_2^* = t + \left[\left(M_{FC}^2\right)^* - \left(M_{HC}^1\right)^* + C_1^*(\lambda_2^*) - C_2^*(\lambda_1^*)\right]/3.$$

The market share of bank 1 is

$$\gamma_1^* = \frac{1}{2} + \frac{1}{6t}(-\left(M_{FC}^1\right)^* + \left(M_{HC}^1\right)^* + C_2^*(\lambda_2^*) - C_1^*(\lambda_1^*)).$$

As $C_1^*(\lambda_1^*)$ and $\left(M_{HC}^1\right)^*$ do not depend on a^W, we have

$$\frac{\partial\gamma_1^*}{\partial a^W} = \frac{1}{6t}\left[-\frac{\partial n_2^*}{\partial a^W}(a^W - c_W) + k\frac{\partial\lambda_2^*}{\partial a^W}F(\lambda_2^*)\right].$$

The profit maximising interchange fee on withdrawals verifies

$$\frac{a^W - c_W}{a^W} = \frac{1}{\epsilon}\frac{1}{1-\frac{2}{3}(2\gamma_1^* - 1)} + \frac{kF(\lambda^*)}{a^W}\frac{\partial\lambda^*/\partial a^W}{\partial n^*/\partial a^W}.$$

Appendix F-3: symmetric issuers as acquirers. A merchant who is affiliated at bank i pays the merchant fee m_i to his bank when the consumers pay by card. He obtains a share α_1 of consumers who are affiliated at bank 1 (who are in proportion γ_1) and a share $1-\alpha_1$ of consumers who are affiliated at bank 2 (who are in proportion $(1-\gamma_1)$). Hence, his transaction costs at stage 4 are

$$\left(C_i^M\right)^* = c_M[\alpha_1\gamma_1 S(\lambda_1)+(1-\alpha_1)(1-\gamma_1)S(\lambda_2)]+m_i[\alpha_1\gamma_1(S-S(\lambda_1))+(1-\alpha_1)(1-\gamma_1)(S-S(\lambda_2))].$$

We have

$$\frac{d\left(C_i^M\right)^*}{dm_i} = \alpha_1\gamma_1(S-S(\lambda_1))+(1-\alpha_1)(1-\gamma_1)(S-S(\lambda_2))+\frac{d\alpha_1}{dm_i}m_i\gamma_1(S-S(\lambda_1))-\frac{d\alpha_1}{dm_i}m_i(1-\gamma_1)(S-S(\lambda_2)).$$

I now express banks' profit. The issuers' profit is different when they are also acquirers. Let me detail here the various differences:

- an issuer receives the interchange fee on card payments if and only if the consumers of the other bank pay by card at one of his affiliated merchants (term $a^C\alpha_i\int_{\lambda_j^*}^{\overline{\lambda}}\frac{F(T)}{T}dT$ in the profit function)
- an issuer has to pay the interchange fee on card payments when its consumers pay by card at one of the merchants that is affiliated at the other bank.
- an issuer receives deposit fees from the merchants (term $\alpha_i M_i$)
- an issuer receives a merchant fee from its affiliated merchants and has to pay the acquisition cost (on the transaction volume $\gamma_i(S-S(\lambda_i))+(1-\gamma_i)(S-S(\lambda_j))$ which corresponds to the total volume of transactions that is paid by card at its affiliated merchant).

Hence, we have

$$\begin{aligned}\pi_i &= \gamma_i A_i(P_i;f_i;w_i)+n_j^*(1-\varphi_j)(a^W-c_W)-k(1-\varphi_j)S(\lambda_j^*)+\\ &\quad a^C\alpha_i\int_{\lambda_j^*}^{\overline{\lambda}}\frac{F(T)}{T}dT+\alpha_i(m_i-c_A)(S-S(\lambda_j))+\alpha_i M_i,\end{aligned}$$

where

$$\begin{aligned}
A_i(P_i; f_i; w_i) &= P_i + (f_i - c_I)\int_{\lambda_i^*}^{\overline{\lambda}} \frac{F(T)}{T}dT + n_i^*((1-\varphi_i)w_i - \varphi_i c_W - (1-\varphi_i)a^W) - a^C(1-\alpha_i)\int_{\lambda_i^*}^{\overline{\lambda}} \frac{F(T)}{T}dT \\
&\quad -k\varphi_i S(\lambda_i^*) + \alpha_i(m_i - c_A)(S - S(\lambda_i)) \\
&\quad -n_j^*(1-\varphi_j)(a^W - c_W) + k(1-\varphi_j)S(\lambda_j^*) - a^C\alpha_i\int_{\lambda_j^*}^{\overline{\lambda}} \frac{F(T)}{T}dT - \alpha_i(m_i - c_A)(S - S(\lambda_j)).
\end{aligned}$$

We solve for the first order conditions by taking the derivative of π_i with respect to m_i, M_i, f_i, w_i and P_i, and we use the fact that we look for a symmetric equilibrium. Using the same reasoning as in Appendix B, I obtain that

$$\begin{aligned}
t_C &= A_i, \\
t_M &= M + a^C\int_{\lambda^*}^{\overline{\lambda}} \frac{F(T)}{T}dT + (m - c_A)(S - S(\lambda)), \\
f &= c_I - a^C - k\varphi\lambda, \\
(1-\varphi)w &= \varphi c_W + (1-\varphi)a^W.
\end{aligned}$$

Notice that there is an infinity of symmetric equilibria, in which banks choose M and m such that $t_M = M + a^C\int_{\lambda^*}^{\overline{\lambda}} \frac{F(T)}{T}dT + (m - c_A)(S - S(\lambda))$. Banks' profit at the equilibrium of stage 2 are:

$$\pi = \frac{t_C}{2} + \frac{t_M}{2} + n^*(1-\varphi)(a^W - c_W) - k(1-\varphi)S(\lambda^*).$$

As the variable part of banks' profit is exactly identical to the case studied in the main model of the article, the profit maximising interchange fees remain the same.

Appendix F-4: ATM Deployment decisions If we consider ATM deployment costs, bank i's profit is given by

$$\pi_i = 2t(\gamma_i^*)^2 + n_j^*(\frac{\rho_i}{\rho_1 + \rho_2})(a^W - c_W) - k(\frac{\rho_i}{\rho_1 + \rho_2})S(\lambda_j^*) - DC(\rho_i).$$

I assume that DC is convex and that it is chosen such that the second-order conditions of profit maximisation are verified. If there is an interior solution, the first-order conditions of profit maximisation with respect to ρ_i are

$$\begin{aligned} & 4t\frac{\partial \gamma_i^*}{\partial \rho_i}\gamma_i^* + \frac{\partial n_j^*}{\partial \rho_i}(\frac{\rho_i}{\rho_1+\rho_2})(a^W - c_W) + n_j^*(\frac{1}{\rho_1+\rho_2})(a^W - c_W) \\ = \quad & k\frac{\partial \lambda_j^*}{\partial \rho_i}F(\lambda_j^*) + k(\frac{1}{\rho_1+\rho_2})S(\lambda_j^*) + DC'(\rho_i). \end{aligned}$$

As the equilibrium is symmetric, we have $\left.\frac{\partial \gamma_i^*}{\partial \rho_i}\right|_{\rho_i=\rho_j=\rho} = 0$, and $\frac{\rho_i}{\rho_1+\rho_2} = \frac{1}{2}$. Hence, banks' investments in ATM deployment satisfy to the following condition:

$$\begin{aligned} & \frac{1}{2}\frac{\partial n^*}{\partial \rho}(a^W - c_W) + \frac{n^*}{2\rho}(a^W - c_W) \\ = \quad & k\frac{\partial \lambda^*}{\partial \rho}F(\lambda^*) + \frac{k}{2\rho}S(\lambda^*) + DC'(\rho). \end{aligned}$$