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

Marianne Verdier*

June 7, 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. If the issuers are ATM owners, I show that the profit maximising interchange fees reflect the trade-off that banks make between the profits on foreign withdrawals and the possibility to save the costs of cash. Social welfare can be increased by regulating interchange fees, and regulatory interventions should be guided by empirical measures of the costs of cash for banks and merchants.

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, Sujit Chakravorti, Nicole Jonker, Wilko Bolt, Emilio Calvano, Jocelyn Donze and Isabelle Dubec for helpful comments and discussions. I also thank several conference participants for their remarks, including the audience of the conference "Economics of Payment IV" at the Federal Reserve Bank of New York (2010).

1 Introduction:

Payment cards are widely held 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 consumers to use their payment cards to withdraw cash. Therefore, payment cards offer consumers a convenient possibility to trade off between cash and card usage at the Point of Sale (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	1.27	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 foreign 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 consumers open accounts. 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

[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). In Europe, a majority of consumers use debit cards when they pay by card.

[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. "These unit costs were calculated at actual average transaction amounts estimated at SEK 165 for cash payments and SEK 620 for card payments".

fees on withdrawals.[5] On the one hand, interchange fees on card payments are 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, if cash withdrawals are not free. 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 the issuers 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, it 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 article 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 network manages a debit card payment system and an ATM network.[7] 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. This model enables me to design some guidelines to determine if interchange fees in debit card systems and in ATM networks are too high or too low to maximise social welfare. The results that I obtain stress the importance of measuring empirically the costs of cash for banks and merchants to improve regulatory decisions.

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

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

[7] In this paper, we do not study the determination of credit card fees. Chakravorti and Emmons (2003), Chakravorti and To (2007) or Bolt and Chakravorti (2008) build models in which the credit function of the card is taken into account.

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 offered 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. 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 debit card, in order to minimize his transaction costs, as in Whitesell (1989). The transaction costs born by a consumer depend on 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 foreign 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 a continuum of monopolies, who differ across their cost of accepting cash. The merchants decide on whether or not to accept debit cards, depending on their cost of cash and on the merchant fee charged by the acquirers, which are perfectly competitive. I assume that the merchants are able to surcharge debit card payments and that there is imperfect pass through of the merchant fee to the consumers.

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. 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 transactions are priced at 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 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.

In 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 the issuers' trade-off between the profits made on foreign withdrawals and the costs of cash. The main contribution of my paper is to show that issuing banks do not necessarily choose high interchange fees on debit card payments when they own ATMs, as this decreases the profit that they make on the ATM side of the market. The mechanism is the following. High interchange fees on card payments encourage consumers to substitute debit cards for cash, provided that the reduction in merchants' acceptance of debit cards is not too high. This change in consumer behavior decreases the profit that banks make on foreign withdrawals, as the demand for withdrawals is reduced, whereas it reduces the costs of cash borne by banks. Therefore, issuing banks choose interchange fees that reflect this trade-off. The fact that banks share their ATMs limits their incentives to choose high interchange fees on card payments, but the interchange fee on card payments is not set to zero, if it is possible to reduce the costs of cash by encouraging consumers to pay by card. The other contribution of my paper is to show that issuing banks tend to choose interchange fees on withdrawals that exceed the monopoly price, because of the negative externalities that they exert on each other when consumers make foreign withdrawals.

The model enables me to design some guidelines for regulatory interventions. The main policy implication of my results is that regulators should not analyze debit card services and cash provision services separately when the payment platform offers both to consumers. If the share of merchants who accept cards is high, social welfare can be increased by raising the interchange fee on withdrawals, provided that the volume of foreign withdrawals is not too high. This measure is likely to encourage consumers to substitute debit cards for cash, and to reduce the costs of merchants who accept cards. If the costs of cash are high for banks and low for merchants, social welfare can be increased by raising the interchange fee on card payments, as this reduces the volume of foreign withdrawals, which increases consumer surplus and banks' profit at a small cost for merchants. If the costs of cash are low for banks and high for merchants, social welfare can be increased by lowering the interchange fee on card payments, provided that the reduction in the volume of transactions paid by debit card is small. This measure is intended to increase merchants' surplus and to encourage merchants to accept cards.

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 theoretical articles.

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. In Whitesell's model, a consumer makes transactions of variable sizes and chooses between paying cash or paying by card according to the value of the transaction. 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 that, in these countries, 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 surcharges on ATMs, which I do not take into account in this article (see Massoud and Bernhardt (2002), or Donze and Dubec (2009)).

[8] 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 E for examples of ATM and payment platforms in various European countries.

[11] Also, there is empirical evidence that consumers trade off between cards and cash according characteristics of the transaction (see Bounie and François (2006), Borzekowski, 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 McAndrews (2003).

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 assuming asymmetric issuers. I also examine the case in which the issuers are also acquirers, and I endogeneize ATM deployment decisions. 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] Each issuing bank offers a debit card to its consumers, which allows them to pay by card at the point of sale (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 maximises banks' joint profit.[14]

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, which enables the consumers to pay by card at the POS, and to withdraw cash from the ATMs. As banks' ATMs are compatible, 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 fixed "foreign fee", w_i, each time they withdraw cash from a foreign ATM, whereas "home" withdrawals are free. I also assume that surcharges 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 flat fee each time they pay by card,

[13] The model will be extended to the case in which the issuers are also acquirers in Section 4.
[14] A summary of all variables is placed in the first Appendix to help the reader.
[15] In the literature on ATMs, a surcharge is a fee that can be imposed by the ATM owner to the foreign consumer that uses the cash dispenser. The foreign fee is a fee that is charged by the issuer of the card to the consumer for conducting a transaction in a foreign ATM.

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. 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]

As in the literature, the merchants' banks, the acquirers, are assumed to be perfectly competitive. The marginal cost of acquiring a debit card transaction is denoted by c_A.

Consumers Consumers are uniformly 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 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),[19] 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 segment, and each consumer is randomly matched to one merchant for each transaction. If $F(T)$ represents the value of spending on all transactions of size T during the

[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 flat fees, while acquiring banks charge proportional fees (which is generally the case in international payment card networks). I do not restrict the card fee to be positive. If the card fee is negative, consumers are granted a reward when they use their debit card.

[17] 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%. In Bergman et al. (2008), it is stated that "ATMs also involve high fixed costs, but there are also substantial variable costs, in particular for filling the machines". In the main model, I neglect the fixed costs of setting up the payment card infrastructure. In the extension section, I will introduce in the model the costs of ATM deployment.

[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' ATM deployment strategies.

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

period, total spending is given by[20]

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

A consumer obtains a surplus V from his purchases, which is assumed to be sufficiently large, such that he always bears 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 debit cards to pay for their expenses:

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

(A2) $\lim\limits_{\lambda \to 0} \dfrac{\lambda}{\sqrt{\int_0^{\lambda} F(T)dT}} = l$, 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. Consumers incur fixed and variable costs and benefits, which differ if they use cash or if they pay by debit card. If a consumer pays by debit card, he has to pay the flat 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.[21] The 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.[22] As there is an opportunity cost $r > 0$, associated to the detention of cash,[23] 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 consumers bear an exogenous fixed cost $b > 0$ when they withdraw cash, which can be interpreted as the time needed to find an ATM.

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

[21] 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 debit card service, but for simplicity, this differentiation is assumed to be exogenous.

[22] 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, 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.

[23] In this model, I assume that r is not a strategic variable that can be decided by the bank which manages the deposit account. The opportunity cost of holding cash is similar to Baumol (1952)'s model of money demand. The Baumol model of money demand is static. For a dynamic model of money demand that takes into account the value of money that is hold at each period for precautionary motives, see Alvarez and Lippi (2009). Also, notice that the Baumol model does not enable me to model the value of each withdrawal.

Given this cost structure, a consumer who has an account at bank i decides to pay by debit 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 $i \in \{1; 2\}$.

Merchants There is a continuum of local monopolies[24] in the economy, who differ across their cost of accepting cash, denoted by c_M,[25] which is distributed over $\left[\underline{c_M}, \overline{c_M}\right]$ according to the probability density h and the cumulative function H, with $\overline{c_M} - \underline{c_M} = 1$. The cost of accepting cash is expressed as a percentage of the transaction value, such that if a consumer pays cash a transaction of size T, a merchant whose cost of accepting cash is c_M bears a cost $c_M T$.[26]

Merchants always accept cash when they do not accept debit cards, as cash is assumed to be legal tender in this economy. Each merchant decides whether or to accept debit cards by comparing the costs of debit cards and cash. For each transaction of value T, each consumer is randomly matched to one merchant, which may refuse cards. When a consumer pays by card, the merchant has to pay a fee to the acquirer, $M(T)$, which is assumed to depend of the size of the transaction, such that $M(T) = mT$, with $m > 0$.[27] The merchants pass through the cost of the merchant fee to the consumers who pay by card at a rate β, in the form of a surcharge, which is assumed to be small.[28] If a consumer pays by card a transaction of value T, he has to pay a surcharge of βmT to the merchant.

[24] 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 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. The welfare implications of this assumption will be discussed in the welfare analysis.

[25] 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. It could also be interpreted as the time needed to collect cash, and count the notes and coins.

[26] It is empirically proved that the cost of accepting cash depends on the value of the transaction. 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.

[27] This assumption is consistent with the industry practices in international networks. In some countries (as the Netherlands for instance), merchants pay flat-fees to accept debit card transactions. I will argue later in the paper that assuming flat fees would not change the mechanisms explained in the article. Also, notice that I do not take into account the fixed costs for merchants of installing POS terminals and setting up cash registers.

[28] The assumption that β is small and exogenous is made for the simplicity of the model, which is already complex. This assumption is consistent with the empirical observations, which tend to show that merchants do not resort frequently to surcharges. If merchants surcharge, the amount of the surcharge is generally small. In this article, a surcharge is defined as a small surcharge if its effect on the consumer behavior is negligible with respect to the effects of the transaction fees. My results would not change if I assumed that merchants are heterogeneous over $c_M/(1-\beta)$, which, as we will see in section 3.3, represents the costs of cash relative to the costs of debit cards for merchants. According to Bolt et al. (2008), surcharging reduces the probability that a consumer pays by debit card by 8%.

Payment system: The payment system chooses the interchange fee on card payments, a^C, and the interchange fee on cash withdrawals, a^W, that 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, and it is assumed to be lower than the marginal cost of the issuer c_I.[29] The interchange fee on cash withdrawals is paid by the issuer of the card to the ATM owner, and is assumed to be higher than the marginal cost c_W.[30] The interchange fees are assumed to be positive and to be paid on a "per-transaction" basis.

I also make the following assumptions:

(A3) For all $i \in \{1, 2\}$, $c_I \leq \overline{\lambda}(v_i - \beta(c_I + c_A) + \sqrt{r(c_W + b)/2S})$.

(A4) $\dfrac{\underline{c_M}}{1-\beta} < c_I + c_A < \dfrac{\overline{c_M}}{1-\beta}$.

(A5) For all $i \in \{1, 2\}$, $\beta(c_I + c_A) < v_i$.

(A6) The level of interchange fees has no impact on the distribution of the goods' prices, F, which does not take into account the surcharges charged by merchants. However, it impacts the transaction costs borne by consumers and merchants.

Assumption (A3) is verified if the variable card benefit, v_i, is high enough. It ensures that 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.[31]

Assumption (A4) is standard in the literature. As we will see in our analysis, it is necessary to ensure that there exists an equilibrium in which consumers do not pay by debit card all their purchases when merchants accept cards.

Assumption (A5) ensures that the surcharge rate is not too high, such that, in equilibrium, the variable benefit of paying by debit card for consumers is always higher than the surcharge than must be paid to the merchant.

Assumption (A6) 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 (including the surcharge rate when they pay by debit card). Empirical studies have

[29] Otherwise, as we will see in Proposition 2, consumers pay for all their expenses by debit card when they meet a merchant who accepts them.

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

[31] 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, if the withdrawals are priced at the marginal cost of the ATM owner and if the interchange fee on card payments is set at the marginal cost of the issuer.

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.[32]

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. Each issuing bank chooses the deposit fee P_i, and the transaction fees (f_i, w_i) for all $i\epsilon\{1,2\}$. The acquirers choose the merchant fee m.
3. Merchants decide whether or not to accept debit cards.
4. Consumers choose the bank from which to hold an account and a payment card.
5. Consumers choose the number of cash withdrawals, n^*, and the threshold which separates debit card and cash payments, λ^*. Then, for each transaction, each consumer is matched randomly to one merchant. If the merchant accepts cards, he pays by debit card if T belongs to $\left[\lambda^*, \overline{\lambda}\right]$ and pays cash otherwise.

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 by the bank in which he has opened an account. He must choose how to pay for a total amount of expenses, S, in order to minimize its transaction costs, as in Whitesell (1989).

[32] See Chang, Evans, and Swartz (2005) for the analysis of the Australian case. Lifting the assumption that the level of interchange fees does not impact retail prices would imply modelling precisely the strategic interactions between merchants and endogenizing the surcharge rate, which is beyond the scope of this paper. This would not suppress the economic mechanism that is explained in this article, but this would also add other effects. More precisely, this would limit the ability of the issuers to use interchange fees on card payments to encourage consumers to pay by card, because of the inflation of retail prices and because of the surcharges that can be imposed by merchants.

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 studying consumers' payment and withdrawal decisions if there is a share α of the merchants who accept cards in the economy. The costs of the payment instruments consist of the fixed costs and the fixed fees paid for each transaction, the variable benefit of paying by card, the opportunity cost of holding cash, and the costs of cash withdrawals. As consumers are randomly matched to merchants, each consumer anticipates that he will be able to choose between paying cash or by card with probability α, and that, with probability $1-\alpha$, he will only be able to pay cash. When it is possible to choose between cash and the debit card, a consumer that holds an account at bank i pays cash if the transaction amount T belongs to $[0, \lambda_i]$ and pays by card if T belongs to $\left[\lambda_i, \overline{\lambda}\right]$. The value of the transactions that are paid cash when the consumer has the choice between cash and the debit card is denoted by $S(\lambda_i)$. From the assumptions on F, we have

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

The expected value of the transactions that are paid cash is given by

$$\alpha S(\lambda_i) + (1-\alpha)S = \alpha \int_0^{\lambda_i} F(T)dT + (1-\alpha)S,$$

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

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

In what follows, I compute the transaction costs borne by a consumer who has opened an account at bank i. Let me start with debit card transactions. There are $F(T)/T$ transactions of size T, and the consumer pays a flat transaction fee f_i when he pays by debit card. A debit card payment occurs with probability α and if T belongs to $\left[\lambda_i, \overline{\lambda}\right]$. As the consumer obtains a variable benefit v_i of paying by card, while paying a surcharge to the merchant, the expected net costs of debit card payments are

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

Finally, the consumer has to bear the costs of withdrawing and holding cash. In average, if $n_i > 0$, the consumer holds a quantity $(\alpha S(\lambda_i) + (1-\alpha)S)/(2n_i)$ of cash in his pocket, so the opportunity cost of cash detention is $r(\alpha S(\lambda_i) + (1-\alpha)S)/(2n_i)$, 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

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

If $n_i > 0$ and λ_i belongs to $[0, \overline{\lambda}]$, 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

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

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, $C_i(\lambda_i, n_i)$. Proposition 1 summarises the results, which are similar to Whitesell (1989). In Proposition 1, I introduce the dummy variable δ, which takes the value 1 if all merchants accept cards, and 0 otherwise.

Proposition 1 *If $f_i > \delta l \sqrt{r((1-\varphi_i)w_i + b)/2}$ 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 debit card if the merchant accepts them. If $f_i \leq \delta l \sqrt{r((1-\varphi_i)w_i + b)/2}$, the consumer pays all his expenses by debit card when he meets a merchant who accepts them. If the card fee is sufficiently high compared to the average cost of using only cash, the consumer never pays by debit card.*

Proof. See Appendix A-1. ■

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

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

n_i^*, is given by

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

The optimal number of withdrawals is expressed as in Baumol (1952)'s model, except that the volume of transactions that is paid cash, $(\alpha S(\lambda_i) + (1-\alpha)S)$, depends on the trade-off that consumers make between cash and the debit card, as in Whitesell (1989). Notice that, in my model, consumers pay all their expenses by card (when it is possible) if the card fee is sufficiently low, as I consider that the consumer behavior depends only on the costs and benefits of the payment instruments.[33]

Some comparative statics.

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. They decrease with the withdrawal fee, with the variable benefit of paying by card, and with the share of merchants who accept cards.*

Proof. See Appendix A-2. ■

When the card fee decreases, a consumer chooses more often to pay by debit 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. If the share of merchants who accept cards increases, there is less uncertainty on the possibility to use the debit card at the point of sale. Consequently, consumers withdraw less cash, and pay more by card.

Now that I have determined the optimal use of payment instruments, I can express the total cost that is born by a consumer as a function of λ_i^*, that is

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

In Lemma 2, I explain how the transaction costs of the consumer vary with the transaction prices charged by the banks, with the costs/benefits of paying by card, and with the share of merchants who accept cards.

[33] In this analysis, I neglect the other attributes of the payment instruments that may be valued by the consumers. For instance, Bolt and Chakravorti (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 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. The consumers' payment costs decrease with the share of merchants who accept cards if the net costs of paying by card are low.*

Proof. See Appendix A-3. ■

Comparison with the switching point that minimizes the costs born by the users. To understand better why consumers make inefficient payment decisions, 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 (the "joint user"). In order to eliminate problems due to merchants' heterogeneity, I assume in the following Lemma that the consumer purchases from one representative merchant, who accepts cards, whose cost of accepting cash is c_M, and who is able to surcharge at a rate $\beta < 1$.

Lemma 3 *Assume that the consumer purchases from one merchant. If $m < c_M/(1-\beta)$, the merchant accepts debit cards, and the consumer's switching point between debit card and cash payments exceeds the level that minimizes the cost of the joint user. If $m = c_M/(1-\beta)$, the consumer's switching point minimizes the cost of the joint user.*

Proof. See Appendix A-4. ■

The first source of inefficiency in payment systems is linked to the fact that the consumer makes the final choice of the payment method without internalizing the impact of this decision on the merchant's costs. As underlined by Bedre and Calvano (2009), several merchants cannot turn down payment cards once they have decided to accept them.[34] Lemma 3 shows that, for each individual merchant, the impact of this inefficiency depends on the cost of cash, on the possibility to surcharge debit card payments, and on the merchant fee. If $m < c_M/(1-\beta)$, the cost of accepting cards for merchants is lower than the cost of accepting cash. As the consumer does not internalize the fact that payment cards are less costly for this merchant, the switching point between card and cash payments exceeds the level that minimizes the total user cost.[35]

[34] Bedre and Calvano (2009) show that this asymmetry in the choice sets of the buyers and the sellers prevents banks from choosing the socially optimal price structure between the consumer side and the merchant side.

[35] "Non strategic" merchants always accept cards when it is less costly for them to do so. However, strategic interactions between merchants may increase "ex post" inefficiencies in the substitution between cash and payment cards, because merchants internalize "ex ante" a part of the consumers' card usage benefit in their decision to accept cards (See Rochet and Tirole (2002)). In this case, the consumers would sometimes choose a switching point between cards and cash that is below the level that minimizes the cost of the joint user - as "ex post", it may be optimal for strategic merchants to refuse cards.

This simple example enables us to understand also why consumers make inefficient payment decisions when merchants are heterogeneous. As perfectly competitive acquirers are generally not able to price discriminate across merchants according to their cost of accepting cash, they charge the same fee to all merchants. Merchants with high values of c_M will tend to accept debit cards, but consumers use them inefficiently, as the switching point is too high to maximise the total user surplus. This effect is reinforced if there is ex ante some uncertainty on the share of merchants who accept cards, because we showed in Lemma 1 that consumers withdraw more cash and pay less by card if they are not sure that cards will be accepted by merchants.

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:

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

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.[36] 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 debit cards. As there are no strategic interactions between merchants, a merchant accepts debit cards if the cost of accepting debit card payments is lower than the cost of cash, that is if

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

[36] If the issuers are asymmetric, none of them corners the market if the differentiation parameter is sufficiently high.

As merchants are heterogeneous over the cost of accepting cash, the percentage of merchants who accept debit cards is defined as $\alpha = P(c_M - m(1-\beta) \geq 0)$. By assumption, the cost of cash c_M is distributed according to the cumulative function H. It follows that the share of merchants who accept cards is given by $\alpha = 1 - H((1-\beta)m)$.[37]

3.4 Stage 2: Bank fees.

In this section, I determine how banks price the transactions and the deposits. Then, I analyse how the prices affect the consumers' payment decisions at stage 5.

3.4.1 The transaction fees and the deposit fee.

At stage 2, banks choose the fees that are charged to the consumers and the merchants. As the acquirers are perfectly competitive, the merchant fee is equal to the acquirer's marginal cost, which is the sum of the interchange fee paid to the issuer of the card and the acquisition cost, that is $m = a^C + c_A$. The share of merchants who accept cards at the equilibrium of stage 2 is given by $\alpha^* = 1 - H((a^C + c_A)(1-\beta))$.

On the market for deposits, 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).[38]

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 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

[37] If merchants pay a flat fee when they accept debit cards, the card acceptance condition is $m \leq c_M T - \beta$, where T is the value of the transaction. If consumers make several transactions with one merchant, merchants decide to accept cards by taking into account the average value of the transaction, as confirmed by empirical studies (see Arango and Taylor (2008)). The mechanism explained in the article is not impacted that the assumption that merchants pay flat fees, as in this case, merchants with high average transaction value (thus high costs of cash) accept cards, while the others refuse cards.

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

to the ATM owner. 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, $k\varphi_i\left(\alpha S(\lambda_i^*)+(1-\alpha)S\right)$. Hence, we have:

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

I now detail the components of the margin that bank i makes on "foreign" consumers' transactions. 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 $k(1-\varphi_j)(\alpha S(\lambda_j^*)+(1-\alpha)S)$. Hence, we have:

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

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 *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 given by*

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

Proof. See Appendix B-1 and B-2. ■

Corollary 1 *At the equilibrium of stage 2, for each bank $i\in\{1;2\}$, the margin on home consumers' transactions is*

$$\left(M_{HC}^i\right)^*=-k\alpha^*\varphi_i\lambda_i^*\int_{\lambda_i^*}^{\bar{\lambda}}\frac{F(T)}{T}dT-k\varphi_i(\alpha^*S(\lambda_i^*)+(1-\alpha^*)S),$$

whereas the margin on foreign consumers' transactions is

$$\left(M_{FC}^i\right)^*=n_j^*(1-\varphi_j)(a^W-c_W)-k(1-\varphi_j)(\alpha S(\lambda_j^*)+(1-\alpha)S).$$

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 use 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 transactions made by "home" consumers are subsidized by the deposit fee, as Corollary 1 shows that the transaction margin on home consumers is negative. Banks charge debit card fees below the marginal cost of a debit card payment to encourage consumers to substitute debit cards for cash. The withdrawal fee, which is charged only for "foreign" withdrawals, reflects the average cost of withdrawals, as $(1-\varphi_i)w_i^* = (1-\varphi_i)a^W + \varphi_i c_W$.[40] This analysis explains why banks often argue that payments are loss-leaders.[41]

When setting the deposit fee, banks take into account the costs and the benefits of attracting consumers. In my model, two effects soften the competition for deposits, because they reduce the benefits of attracting "home" consumers. The first effect is the possibility to make profit on "foreign" consumers through foreign withdrawals. Hence, a consumer pays in the deposit fee the opportunity cost for bank i of losing a "foreign" consumer when it attracts a depositor (term $\left(M_{FC}^i\right)^*$ in the deposit fee). This effect is also present in Massoud and Bernhardt (2002), or Donze and Dubec (2006). The second effect that softens the competition for deposits is related to the variable costs of cash generated by the consumers who withdraw a lot from the bank in which they hold an account. This effect, which has not been modelled in the literature, 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. When they attract a consumer, banks take into account the fact that their margin on the consumer's transaction will be negative, because of the costs of cash.

[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] In section 4, I endogeneize φ_i, by assuming that it depends on banks' investments in ATM deployment. It would be also possible to assume that φ_i increases with the level of the foreign fee w_i. In this case, the foreign fee is a non linear tariff, which makes it much more difficult to solve the problem numerically. The foreign fee depends on the elasticity of the demand for foreign withdrawals with respect to the foreign fee.

[41] According to Van Hove (2008), 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 (2008) 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". According to Bergman et al. (2008) the card market is profitable for Swedish banks (SEK 460 million) but banks incur a loss on cash distribution (of SEK 466 million). They say that "it may thus be concluded that cash distribution is being subsidized by profits made in the card market".

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

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

In the following section of the paper, I will focus on the case in which banks are perfectly symmetric,[42] and I will denote their joint profit in this case by π, where

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

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

In this section, 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 on 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.

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 some 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. ■

Consumers use both cash and cards to pay for their expenses if the interchange fee on card payments is not too high. Lemma 5 explains how interchange fees impact consumers' decision to substitute debit cards for cash.

Lemma 5 *If the interchange fee on withdrawals increases, the consumers choose more often to pay by debit card, and make fewer withdrawals. The total transaction costs born by the*

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

consumers increase.

If β is sufficiently small and if the sensitivity of α^* to a^C is small (resp. high), the threshold above which consumers pay by debit card decreases (resp. increases) when the interchange fee on card payments increases, and the consumers make fewer (resp. more) withdrawals. The total transaction costs borne by the consumers decrease.

Proof. See Appendix C-2. ■

An increase in the interchange fee on withdrawals raises the average marginal cost of each ATM owner, if consumers make some foreign withdrawals. Consequently, the withdrawal fee becomes higher and consumers reduce their volume of cash payments. It follows that high interchange fees on withdrawals encourage consumers to substitute debit card for cash.[43]

An increase in the interchange fee on card payments has three effects on consumers' incentives to substitute debit cards for cash. The first effect of an increase in a^C is to reduce the perceived marginal cost of the issuers. As the issuers price the transactions at the average cost, the card fee becomes lower, which, from Lemma 1, encourages consumers to pay by card more frequently. The second effect is related to the presence of surcharges in consumers' transaction costs. A rise in a^C increases the merchant fee, which raises the amount of surcharges paid by consumers. This effect is small, as we assumed that the effect of surcharges on consumer behavior can be neglected in a first approach. The third effect is linked to the fact that consumers anticipate the share of merchants who accept cards in their payments decisions. If the interchange fee on card payments increases, the share of merchants who accept cards becomes lower. This third effect has a negative impact on consumers' incentives to pay by debit card, as debit cards are less likely to be accepted by merchants. If the third effect is small compared to the first effect, a rise in the interchange fee on card payments increases consumers' incentives to substitute debit cards for cash. Otherwise, if the interchange fee on card payments has a strong impact on the share of merchants who accept cards, a rise in the interchange fee on card payments may reduce consumers' incentives to pay by card, because of higher uncertainty about card acceptance.

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 max-

[43] Notice that this effect exists even if all withdrawals are free for consumers, that is if w_i is constrained to be equal to zero. In this case, the card fee is negatively related to the interchange fee on withdrawals, as banks subsidize card payments to encourage consumers to substitute cards for costly foreign withdrawals.

imisation, that is, from (7),

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

and

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

Notice that, if there are no foreign withdrawals, that is if $\varphi = 1$, the level of interchange fees has no impact on banks' joint profit. Proposition 3 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,*

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

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,

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

If there are no costs of cash for banks or if consumers pay less by card when the interchange fee on card payments increases, the profit maximising interchange fee on card payments is equal to zero.

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 costs of cash. The profit maximising interchange fee on withdrawals is set by the platform such that the markup of the ATM owner 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 costs of cash for banks. The higher the costs of cash for banks, the higher the interchange fee on withdrawals.

As noted in Lemma 5, a rise in the interchange fee on card payments provides consumers with incentives to substitute debit cards for cash if the sensitivity of the share of merchants who accept cards to a^C is small. This change in consumers' behavior has two effects on banks' joint profit. First, this decreases the profit made on foreign withdrawals, as the number of foreign withdrawals becomes lower (term $(a^W - c_W)\partial n^*/\partial a^C$ in (9)). Second, the costs of cash are reduced, as consumers withdraw less cash from the ATM network (term $(a^W - c_W)\partial \lambda^*/\partial a^C$ in (9)). In addition to these effects, a rise in the interchange fee on card payments decreases merchants' acceptance of debit cards, which increases the costs of cash borne by banks (term $k\dfrac{\partial \alpha^*}{\partial a^C}(S - S(\lambda^*))$ in (9)). This last effect may prevent issuing banks from choosing excessive interchange fees on card payments, as they generate low levels of substitution of debit cards for cash if card acceptance is reduced.

In Lemma 5, we also noted that consumers pay less by card when the interchange fee on card payments increases, if the sensitivity of α^* to a^C is high. In this case, the profit maximising interchange fee on card payments is equal to zero, as positive interchange fees on card payments cannot contribute to decrease the costs of cash, whereas they reduce the profits made on foreign withdrawals.

Let me compare the result of Proposition 3 with the existing literature on payment cards. In Rochet and Tirole (2002) or Wright (2004), the merchants' resistance to card acceptance - or the merchants' elasticity of demand if merchants are heterogeneous- is the only counterbalance that limits the issuers' ability to choose a high interchange fee. By modelling explicitly the costs of cash for issuing banks, and the profits made on foreign withdrawals, my paper shows that issuing banks which are also ATM owners choose an interchange fee on card payment 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.[44]

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 profit 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

[44] In this article, the profits made on the card side depend only on "costs savings". However, banks have other motivations for selling debit cards, such as the possibility to extract consumer surplus by selling additional services, or by increasing the differentiation from their competitors.

payment instrument, that is,

$$MS = S-(1-\alpha^*)E(c_M/c_M \leq c_M^*)S-(1-\beta)\alpha^*(a^C+c_A)(S-S(\lambda^*))-\alpha^*S(\lambda^*)E(c_M/c_M \geq c_M^*), \tag{10}$$

where E denotes the mathematical expectation.

A rise in the interchange fee on card payments increases the cost of accepting debit cards compared to the cost of accepting cash. Hence, fewer merchants accept debit cards and the share of merchants who accept debit cards decreases. Consumers also react to changes in the interchange fee on card payments. From Lemma 5, if the sensitivity of α^* to a^C is not too high, the increase in a^C reduces the total value of cash payments and increases the value of debit card payments, as consumers decide to increase the substitution of debit cards for cash. This change in consumer behavior has no impact on the merchants who accept only cash, whereas it decreases the costs borne by the merchants who accept debit cards. This is because debit cards remain less costly for merchants who accept them, even if the merchant fee has increased. The overall effect on an increase in a^C on the merchant surplus is ambiguous.[45]

An increase in the interchange fee on withdrawals has no impact on the cost borne by merchants for accepting payment instruments. But it changes consumer behavior, as consumers withdraw less cash and pay more by debit card when foreign withdrawals become more costly. This effect increases the surplus of merchants who accept cards, as these merchants prefer to be paid by debit card, whereas the surplus of merchants who accept only cash is unchanged. 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.[46]

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

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

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

[46] The merchants' surplus is not reduced when the interchange fees on withdrawals increase, if consumers' demand does not decrease when transactions become more costly. This is the case in my model, as the volume of spendings is constant by assumption.

where from Proposition 2 and Corollary 1,

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

and from (1), to

$$C^* = \sqrt{2r((1-\varphi)a^W + \varphi c_W + b)\left(\alpha^* S(\lambda^*) + (1-\alpha^*)S\right)} \tag{13}$$

$$+\alpha^*(c_I - a^C - k\varphi\lambda^*) \int_{\lambda^*}^{\overline{\lambda}} \frac{F(T)}{T} dT + \alpha^*(\beta(a^C + c_A) - v)(S - S(\lambda^*)). \tag{14}$$

To understand how variations in interchange fees impact consumer surplus, we have to analyze the impact of interchange fee variations on the transaction costs borne by consumers and on the deposit fee.

A higher a^W increases the costs of foreign withdrawals for consumers, which raises their transaction costs. A lower demand for withdrawals has two different effects on the competition for deposits. First, the costs of cash for banks are reduced, which increases the benefits of attracting home consumers. This effect toughens the competition for deposits. Second, the margin on foreign withdrawals may either increase or decrease, as the volume of foreign withdrawals is lower, but the margin per foreign withdrawal is higher. If the margin on foreign withdrawals increases, this effect softens the competition for deposits. To sum up, it is impossible to sign the impact of an increase in the interchange fee on withdrawals on consumer surplus.

A higher a^C decreases the transaction costs borne by consumers if the sensitivity of α^* to a^C is small. An increase in a^C has two opposite effects on the deposit fee. As consumers pay more by card, the costs of cash for banks are reduced, which softens the competition for deposits and increase the deposit fee. However, as the volume of foreign withdrawals is smaller, the margin on foreign withdrawals decreases, which toughens the competition for deposits. In general, it is also impossible to sign the variation of the consumer surplus with the interchange fee on card payments (See Appendix D-2).

Social welfare improvement. In Proposition 4, I explain how social welfare can be increased by changing the level of one interchange fee, while keeping the other interchange fee constant. Social welfare is defined as the sum of consumer surplus, merchant surplus and bank surplus. I denote by $\left((a^C)^\pi, (a^W)^\pi\right)$ the pair of interchange fees that maximises banks' joint profit and by $\left((a^C)^{SW}, (a^W)^{SW}\right)$ the pair of interchange fees that maximises social welfare.

Proposition 4 *If the share of merchants who accept cards is high (resp. small) at* $\left((a^C)^\pi,(a^W)^\pi\right)$*, social welfare can be increased by raising (resp. lowering) the interchange fee on withdrawals, provided that the volume of foreign withdrawals is not too high (resp. not too low) at* $\left((a^C)^\pi,(a^W)^\pi\right)$*.*
If the costs of cash are high for banks and low for merchants, social welfare can be increased by raising the interchange fee on card payments.
If the costs of cash are low for banks and high for merchants, social welfare can be increased by lowering the interchange fee on card payments, provided that the decrease in the volume of transactions paid by debit card is small.

Proof. See Appendix D-3. ■

The pair of interchange fees that maximises banks' joint profit does not maximise social welfare. This is because issuing banks - which are not acquirers- do not internalize the effect of their price decisions on merchants' costs of accepting debit cards. Also, issuing banks who are ATM owners can exert their market power on consumers by choosing interchange fees on withdrawals that exceed the monopoly price.

Measuring the costs of cash for the agents involved in the payment process (banks, consumers, merchants) is an essential step to guide regulatory intervention. In what follows, I provide some intuitions of the policy recommendations given in Proposition 4, which are summarized in a table placed in Appendix D-3. If the share of merchants who accept cards is high at the profit maximising interchange fees, social welfare can be increased by raising the interchange fee on withdrawals. This measure is likely to encourage consumers to substitute debit cards for cash, and to lower the costs of the merchants who accept debit cards, as a greater volume of transactions will be paid by card. Respectively, if the share of merchants who accept cards is low, social welfare can be increased by lowering the interchange fee on withdrawals, as this will decrease the transaction costs borne by consumers, who use cash intensively.

If the costs of cash are high for banks and low for merchants,[47] social welfare can be increased by raising the interchange fee on card payments for a given level of the interchange fee on withdrawals. This measure is going to encourage consumers to pay by debit card, if the share of merchants who accept cards does not decrease too much, and to increase the volume of transactions paid by debit card, which is going to benefit banks.[48]

[47] If the costs of cash are low for merchants, the share of merchants who accept cards at the equilibrium of stage 2, α^*, is small.

[48] In this case, however, the regulator should keep in mind that strategic interactions between merchants could offset this result. Strategic merchants are ready to accept debit cards even if the cost of debit cards is higher than the cost of cash. Hence, if the best policy intervention is to raise the interchange fee on card payments, this could generate a higher fall in merchant surplus when merchants are strategic than in our situation.

If the costs of cash are low for banks and high for merchants, social welfare can be increased by lowering the interchange fee on card payments for a given level of the interchange fee on withdrawals. This measure is likely to increase card acceptance, which may raise the volume of transactions paid by debit card, and eventually reduce the costs of cash for banks. This is also going to increase merchant surplus by lowering the costs of the merchants who accept cards.

4 Extensions and discussions.

In this section, I provide some extensions of the results obtained in section 3. First, I consider the case of asymmetric issuers. Second, I study the case in which symmetric issuers are also acquirers. Finally, I discuss the case of endogeneous ATM deployment decisions.

4.1 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 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 their 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 depositors, by internalizing the effects of 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, consider 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

$$\pi = 2t(\gamma_1^*)^2 + 2t(1-\gamma_1^*)^2 + n_2^*(a^W - c_W) - k\left[\alpha^* S(\lambda_2^*) + (1-\alpha^*)S\right],$$

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. The main modification of Proposition 3 stems from the fact that the issuers' interests are not aligned (See Appendix F-1). 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). Regulators should take into account banks' market shares in their decision to intervene in debit card markets, as regulatory decisions may have a positive effect on some banks to the detriments of the others.

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. I have suggested remedies to cope with inefficiencies that arise when there is inefficient substitution between debit card and cash for a given market structure. But I have not studied another potential solution which is to promote entry on the market.

4.2 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 $\beta = 0$ and that merchants are homogenous as regards to their costs of accepting payment instruments. The merchants are uniformly located along the linear city, and may travel to open an account either at bank 1 or at bank 2. 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 $t_M\,|x - d_i| + M_i + \left(C_i^M\right)^*$, and obtains a surplus, which I assume to be sufficiently large such that the market is covered.[49] 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

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

[49] We have studied a simple case in which all merchants accept cards at the equilibrium. If merchants are heterogeneous over their cost of accepting cash, the results would be similar to Proposition 3 in a symmetric equilibrium, except that now, banks compete on two markets: the market of the merchants who refuse cards and the market of the merchants who accept them.

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 (if $\alpha^* = 1$), 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. A bank that has a high market share on the acquisition side will also prefer a low interchange fee on withdrawals, because its profit becomes higher when consumers increase the substitution between cards and cash.

4.3 Endogenous ATM deployment decisions.

In the model that is developped in 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.

For this purpose, 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 $\varphi_i = \rho_i/(\rho_1 + \rho_2)$.[50] I also assume that the fixed cost of withdrawing cash, b, is decreasing with the total number of ATMs in the economy, such that investments in ATM deployment reduce the costs of making a withdrawal for a consumer. 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 that is presented in section 3).

In Appendix F-3, 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

[50] The uniform probability assumption is biased if home withdrawals are less costly than foreign withdrawals, because consumers tend to withdraw more from their bank when home withdrawals are free. But it is useful to understand how endogenous ATM deployment decisions impact the results obtained in this article. Taking into account the prices would make it difficult to solve the model numerically.

benefit of investments in ATM deployment is equal to the marginal cost, that is

$$\frac{1}{2}\frac{\partial n^*}{\partial \rho}(a^W - c_W) + \frac{n^*}{4\rho}(a^W - c_W) = \frac{k\alpha^*}{2}\frac{\partial \lambda^*}{\partial \rho}F(\lambda^*) + \frac{k}{4\rho}\left[\alpha^* S(\lambda^*) + (1-\alpha^*)S\right] + DC'(\rho). \quad (15)$$

The marginal benefit of investing in ATM deployment is related 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 when they make their investment decision (term $\frac{k\alpha^*}{2}\frac{\partial \lambda^*}{\partial \rho}F(\lambda^*) + \frac{k}{4\rho}\left[\alpha^* S(\lambda^*) + (1-\alpha^*)S\right]$ in (15)). 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.

Assuming 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. Social welfare can be increased by regulating interchange fees on withdrawals or interchange fees on card payments, and the optimal regulatory intervention depends on the costs of cash borne by banks and merchants respectively.

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. Or that banks can subsidize intensively debit card fees to reduce the costs of cash. 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.[51] 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.[52] 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] ALVAREZ, F. & LIPPI, F. (2009): "Financial Innovation and the Transaction Demand for Cash", Econometrica, Econometric Society, vol. 77 (2), pp. 363-402.

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

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

[51] See DNB Quaterly Bulletin, March 2006.

[52] 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".

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

[5] 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.

[6] BEDRE O. & CALVANO E. (2009): "Pricing Payment Cards", ECB Working Paper No 119.

[7] 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.

[8] BERGMAN M., GUIBOURG G. & SEGENDORF B. (2008): "Card and Cash payments – the social perspective in Sweden" in Chapter 2 of "Evolving Payment Habits. Proceedings of the Bank of Finland payment habits seminar 2008".

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

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

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

[12] 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.

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

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

[15] BOLT W., JONKER N. & VAN RESENLAAR C. :"An empirical analysis of payment behavior and debit card surcharges in the Netherlands" in Chapter 7 of "Evolving Payment Habits. Proceedings of the Bank of Finland payment habits seminar 2008".

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

[17] 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.

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

[19] 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.

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

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

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

[23] 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.

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

[25] 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.

[26] 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.

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

[28] 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.

[29] 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.

[30] 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.

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

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

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

[34] 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.

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

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

[37] 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.

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

[39] VAN HOVE L. (2008): "On the War on Cash and its Spoils", International Journal of Electronic Banking, vol. 1, no. 1, pp. 36-45.

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

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

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

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

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

[45] 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 0: Summary of all variables In the following table, we summarize the variables used in the model:

	Exogenous variables:	Endogenous variables
Consumers	φ_i percentage of home withdrawals	λ_i threshold value cash/card
	b fixed cost of withdrawals	n_i number of withdrawals
	v_i variable benefit of paying by card	C_i transaction costs
	r opportunity cost of cash detention	
	S total value of spendings	
	$F(T)$ value of spendings of size T	
Merchants	c_M cost of cash	α % of merchants who accept cards
	β cost of cash	
	H distribution of the cost of cash	
Banks	c_I issuing cost	P_i deposit fee
	c_A acquisition cost	w_i withdrawal fee
	c_W cost of a withdrawal	f_i debit card fee
	k variable cost of cash	m merchant fee
		γ_i market share
Platform	a^C interchange fee on card payment	
	a^W interchange fee on withdrawals	

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 transaction costs. I start by determining the equations verified by an interior optimum, if $\lambda_i > 0$, $n_i > 0$ and $\alpha > 0$. Then I derive the conditions that are necessary and sufficient for this optimum to exist. Solving for the first order conditions of profit maximisation,[53] I obtain

$$\frac{\partial C_i(\lambda_i, n_i)}{\partial n_i} = (1-\varphi_i)w_i + b - \frac{r\left(\alpha S(\lambda_i) + (1-\alpha)S\right)}{2(n_i)^2} = 0, \tag{16}$$

and

$$\frac{\partial C_i(\lambda_i, n_i)}{\partial \lambda_i} = \frac{r\alpha}{2n_i}\frac{\partial S(\lambda_i)}{\partial \lambda_i} - \alpha f_i \frac{F(\lambda_i)}{\lambda_i} + \alpha(v_i - \beta m)\frac{\partial S(\lambda_i)}{\partial \lambda_i} = 0. \tag{17}$$

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 (16), I obtain the optimal number of withdrawals,

$$n_i^* = \sqrt{\frac{r\left(\alpha S(\lambda_i) + (1-\alpha)S\right)}{2((1-\varphi_i)w_i + b)}}. \tag{18}$$

As $S(\lambda_i) = \int_0^{\lambda_i} F(T)dT$, we have $\partial S(\lambda_i)/\partial \lambda_i = F(\lambda_i)$. Replacing for n_i^* in (17), if there is an interior solution, λ_i^*, then it must satisfy

$$\alpha F(\lambda_i^*)\left[\sqrt{\frac{r((1-\varphi_i)w_i + b)}{2\left(\alpha S(\lambda_i^*) + (1-\alpha)S\right)}} - \frac{f_i}{\lambda_i^*} + (v_i - \beta m)\right] = 0, \tag{19}$$

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

$$\lambda_i^*\sqrt{\frac{r((1-\varphi_i)w_i + b)}{2\left(\alpha S(\lambda_i^*) + (1-\alpha)S\right)}} - f_i + (v_i - \beta m)\lambda_i^* = 0. \tag{20}$$

I now show that, under some conditions, there exists a unique λ_i^* that verifies equation (20). For this purpose, let

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

My aim is to derive the conditions under which there exists a unique $\lambda_i^* \in \left(0, \overline{\lambda}\right)$ such that

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

$g(\lambda_i^*) = 0$. The function g is differentiable over $\left(0;\overline{\lambda}\right]$, and its first derivative is

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

Let the function μ be defined as $\mu(\lambda) = 2\left(\alpha S(\lambda) + (1-\alpha)S\right) - \lambda\alpha F(\lambda)$ over $\left[0,\overline{\lambda}\right]$. We have

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

I now study the function μ over $\left[0,\overline{\lambda}\right]$. As $S^{'}(\lambda) = F(\lambda)$, then $\mu^{'}(\lambda) = \alpha F(\lambda) - \lambda\alpha F^{'}(\lambda)$, and $\mu^{"}(\lambda) = -\lambda\alpha F^{"}(\lambda)$. As F is concave by assumption (A1), μ is convex. So $\mu^{'}$ is increasing over $\left[0,\overline{\lambda}\right]$. As $\mu^{'}(0) = \alpha F(0) \geq 0$, $\mu^{'}$ is positive over $\left[0,\overline{\lambda}\right]$. Hence, μ is increasing over $\left[0;\overline{\lambda}\right]$. As $\mu(0) = 2(1-\alpha)S \geq 0$, μ is positive over $\left[0,\overline{\lambda}\right]$. Hence, as $v_i - \beta m > 0$ if β is small,[54] the function $\mu^{'}$ is strictly positive over $\left(0,\overline{\lambda}\right]$ and consequently, μ is increasing over $\left(0,\overline{\lambda}\right]$.

By assumption (A3), if $\alpha = 1$, we have

$$\lim_{\lambda \longrightarrow 0}(\lambda/\sqrt{\alpha S(\lambda) + (1-\alpha)S}) = l.$$

Otherwise, if $0 \leq \alpha < 1$, we have

$$\lim_{\lambda \longrightarrow 0}(\lambda/\sqrt{\alpha S(\lambda) + (1-\alpha)S}) = 0.$$

We introduce the dummy variable δ which takes the value 1 if $\alpha = 1$, and the value 0 otherwise. The function g can be prolongated to $g(0) = \delta l\sqrt{r((1-\varphi_i)w_i + b)/2} - f_i$. If $f_i > \delta l\sqrt{r((1-\varphi_i)w_i + b)/2}$, we have $g(0) < 0$. If $g(0) < 0$ 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 $f_i < \delta l\sqrt{r((1-\varphi_i)w_i + b)/2}$, we have $g(0) > 0$. If $g(0) > 0$ 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

$$f_i < (v_i - \beta m)\overline{\lambda} + \overline{\lambda}\sqrt{r((1-\varphi_i)w_i + b)/2S}. \qquad \text{(Condition (A-1))}$$

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.

[54] This condition will be satisfied in equilibrium because of Assumption (A5). See Appendix C.

If $f_i \leq \delta l \sqrt{r((1-\varphi_i)w_i+b)/2}$, from (17) 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 all his expenses by debit card when it is possible. If the card fee is higher than the average cost of paying everything cash, that is if $g(\overline{\lambda}) < 0$, the consumer never pays by debit 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: Proof of Lemma 1. I now prove that λ_i^* and n_i^* increase with f_i and m, and that they decrease with w_i, v_i, and α. I start by showing that λ_i^* and n_i^* increase with f_i. Taking the derivative of (20) with respect to f_i, I obtain that

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

that is, after some rearrangements,

$$\frac{\partial \lambda_i^*}{\partial f_i}\left(v_i - \beta m + \sqrt{\frac{r((1-\varphi_i)w_i+b)}{2(\alpha S(\lambda_i^*)+(1-\alpha)S)}}\frac{\mu(\lambda_i^*)}{2(\alpha S(\lambda_i^*)+(1-\alpha)S)}\right) = 1, \tag{21}$$

where $\mu(\lambda) = 2\left(\alpha S(\lambda) + (1-\alpha)S\right) - \lambda \alpha F(\lambda)$. I have already shown in Appendix A-1 that μ is positive over $\left[0, \overline{\lambda}\right]$. Therefore, all the terms in the parenthesis of (21) 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 (18) with respect to f_i, I find that

$$\frac{\partial n_i^*}{\partial f_i} = \frac{\alpha F(\lambda_i^*)n_i^*}{2(\alpha S(\lambda_i^*)+(1-\alpha)S)}\frac{\partial \lambda_i^*}{\partial f_i}. \tag{22}$$

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 (20) with respect to w_i, I obtain that

$$\begin{aligned}
&\frac{\partial \lambda_i^*}{\partial w_i}\left(v_i - \beta m + \sqrt{\frac{r((1-\varphi_i)w_i+b)}{2(\alpha S(\lambda_i^*)+(1-\alpha)S)}}\left(\frac{\mu(\lambda_i^*)}{2(\alpha S(\lambda_i^*)+(1-\alpha)S)}\right)\right) \\
= \quad & -\lambda_i^*\sqrt{\frac{r((1-\varphi_i)w_i+b)}{2(\alpha S(\lambda_i^*)+(1-\alpha)S)}}\frac{1}{2((1-\varphi_i)w_i+b)}.
\end{aligned}$$

The expression in the parenthesis of the left side of this equality is positive if β is sufficiently small as in (21). The right side of the equation is negative. It follows that $\partial \lambda_i^*/\partial w_i \leq 0$. If the price of foreign withdrawals rises, this reduces the threshold above which consumers pay by card.

Taking the derivative of (18) with respect to w_i, I obtain that

$$\frac{\partial n_i^*}{\partial w_i} = \frac{\alpha F(\lambda_i^*) n_i^*}{2(\alpha S(\lambda_i^*) + (1-\alpha)S)} \frac{\partial \lambda_i^*}{\partial w_i} - \frac{(1-\varphi_i) n_i^*}{2((1-\varphi_i) w_i + b)}, \tag{23}$$

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.

Using a similar proof, we can prove that λ_i^* and n_i^* increase with m, and that they decrease with v_i and α.

Appendix A-3: Proof of Lemma 2. Using the envelop's theorem, from (1), I find that

$$\left. \frac{\partial C_i^*(\lambda_i, n_i, w_i, f_i, \alpha, m, v_i)}{\partial w_i} \right|_{(\lambda_i^*, n_i^*)} = (1-\varphi_i) n_i^* \geq 0. \tag{24}$$

Using the same reasoning, I obtain that

$$\left. \frac{\partial C_i^*(\lambda_i, n_i, w_i, f_i, \alpha, m, v_i)}{\partial f_i} \right|_{(\lambda_i^*, n_i^*)} = \alpha \int_{\lambda_i^*}^{\bar{\lambda}} \frac{F(T)}{T} dT \leq 0. \tag{25}$$

Similarly, I have that

$$\left. \frac{\partial C_i^*(\lambda_i, n_i, w_i, f_i, \alpha, m, v_i)}{\partial v_i} \right|_{(\lambda_i^*, n_i^*)} = -\alpha (S - S(\lambda_i^*)) \leq 0, \tag{26}$$

and

$$\left. \frac{\partial C_i^*(\lambda_i, n_i, w_i, f_i, \alpha, m, v_i)}{\partial m} \right|_{(\lambda_i^*, n_i^*)} = \beta \alpha (S - S(\lambda_i^*)) \geq 0. \tag{27}$$

Finally,

$$\left. \frac{\partial C_i^*(\lambda_i^*)}{\partial \alpha} \right|_{(\lambda_i^*, n_i^*)} = \frac{S(\lambda_i^*) - S}{2} \sqrt{\frac{r((1-\varphi_i) w_i + b)}{2(\alpha S(\lambda_i^*) + (1-\alpha) S)}} + f_i \int_{\lambda_i^*}^{\bar{\lambda}} \frac{F(T)}{T} dT + (\beta m - v_i)(S - S(\lambda_i^*)). \tag{28}$$

As $S - S(\lambda_i^*) \geq 0$ and $\beta m - v_i \leq 0$ if β is sufficiently small, equation (28) is negative if f_i is sufficiently small, or if v_i is high.

So the consumer's total cost increases with the withdrawal fee, the card fee and the merchant fee, whereas it decreases with the variable benefit of paying by card. It decreases with the share of merchants who accept cards if the card fee is small or if the variable benefit of paying by card is high.

Appendix A-4: comparison with the joint cost minimizing switching point. Assume that the consumer purchases from one representative merchant who accepts cards. 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, that is

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

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 debit card verifies

$$F((\lambda_i^*)_{TU})\left[\frac{r}{2(n_i^*)_{TU}} - \frac{f_i}{(\lambda_i^*)_{TU}} + (v_i - m + c_M)\right] = 0. \tag{29}$$

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 that minimizes the costs of the consumer. We use the function g defined in Appendix A-1, that is

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

From Appendix A-1, the threshold χ chosen by the consumer solves $g(\chi) = 0$. From (29), the threshold above which the "joint user" would like to pay by debit card verifies

$$g((\lambda_i^*)_{TU}) = (\lambda_i^*)_{TU}\left((1-\beta)m - c_M\right).$$

As the merchant accepts debit cards, the cost of cash must be higher than the cost of debit cards, that is $(1-\beta)m - c_M \leq 0$. It follows that $g((\lambda_i^*)_{TU}) \leq g(\chi)$. We proved in Appendix A-1 that g is increasing over $\left[0, \overline{\lambda}\right]$. Hence, we can conclude that $(\lambda_i^*)_{TU} \leq \chi$.

Appendix B: Proof of proposition 2. I assume that t, the differenciation parameter, is sufficiently high to ensure 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 (See Appendix C).

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

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

where

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

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

$$\frac{\partial \pi_i}{\partial P_i} = \frac{\partial \gamma_i}{\partial P_i} A_i(P_i; f_i; w_i) + \gamma_i \frac{\partial A_i(P_i; f_i; w_i)}{\partial P_i} = 0, \tag{30}$$

$$\frac{\partial \pi_i}{\partial f_i} = \frac{\partial \gamma_i}{\partial f_i} A_i(P_i; f_i; w_i) + \gamma_i \frac{\partial A_i(P_i; f_i; w_i)}{\partial f_i} = 0, \tag{31}$$

and

$$\frac{\partial \pi_i}{\partial w_i} = \frac{\partial \gamma_i}{\partial w_i} A(P_i; f_i; w_i) + \gamma_i \frac{\partial A(P_i; f_i; w_i)}{\partial w_i} = 0. \tag{32}$$

I start by equation (30). 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 (30), I obtain that

$$A(P_i; f_i; w_i) - 2t\gamma_i = 0. \tag{33}$$

Now consider equation (31). From (2), I find that $\partial\gamma_i/\partial f_i = -(1/2t)\partial C_i^*(\lambda_i^*)/\partial f_i$. Replacing for this expression in (31), using (33), and after a simplification by $\gamma_i > 0$, equation (31) can be rewritten as

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

As

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

and as from (25) $\frac{\partial C_i^*(\lambda_i^*)}{\partial f_i}=\alpha\int_{\lambda_i^*}^{\overline{\lambda}}\frac{F(T)}{T}dT$, equation (31) becomes

$$\frac{\partial\lambda_i^*}{\partial f_i}\frac{F(\lambda_i^*)}{\lambda_i^*}\alpha(f_i+a^C-c_I)-\frac{\partial n_i^*}{\partial f_i}\left[(1-\varphi_i)(w_i-a^W)-\varphi_i c_W\right]+k\alpha\varphi_i F(\lambda_i^*)\frac{\partial\lambda_i^*}{\partial f_i}=0. \qquad (34)$$

From (18), I obtain that

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

Replacing for this expression in (34), I get the equation that defines the card fee

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

Therefore,

$$f_i+a^C-c_I-\frac{n_i^*\lambda_i^*}{2}\frac{((1-\varphi_i)(w_i-a^W)-\varphi_i c_W)}{(\alpha S(\lambda_i^*)+(1-\alpha)S)}+k\varphi_i\lambda_i^*=0. \qquad (35)$$

I now study equation (32). From (2), I obtain $\partial\gamma_i/\partial w_i=-(1/2t)\partial C_i^*(\lambda_i^*)/\partial w_i$. Hence, replacing for this expression in (32), after a simplification by $\gamma_i>0$, we find that equation (32) is equivalent to

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

Using (24), I find that $\partial C_i/\partial w_i=(1-\varphi_i)n_i^*$. As

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

replacing in (32) and using (33), equation (32) can be rewritten as

$$\frac{\partial\lambda_i^*}{\partial w_i}\frac{\alpha F(\lambda_i^*)}{\lambda_i^*}(f_i+a^C-c_I)-\frac{\partial n_i^*}{\partial w_i}((1-\varphi_i)(w_i-a^W)-\varphi_i c_W)+k\alpha\varphi_i F(\lambda_i^*)\frac{\partial\lambda_i^*}{\partial w_i}=0. \qquad (36)$$

From (18), I have that $\frac{\partial n_i^*}{\partial w_i}=\frac{\alpha n_i^* F(\lambda_i^*)}{2(\alpha S(\lambda_i^*)+(1-\alpha)S)}\frac{\partial\lambda_i^*}{\partial w_i}-\frac{(1-\varphi_i)n_i^*}{2((1-\varphi_i)w_i+b)}$. Replacing in

(36), I obtain that

$$-\frac{\partial \lambda_i^*}{\partial w_i}\frac{\alpha F(\lambda_i^*)}{\lambda_i^*}\left[f_i + a^C - c_I - \frac{n_i^*\lambda_i^*}{2(\alpha S(\lambda_i^*)+(1-\alpha)S)}((1-\varphi_i)(w_i - a^W) - \varphi_i c_W) + k\varphi_i\lambda_i^*\right] \quad (37)$$

$$= \frac{-n_i^*\lambda_i^*((1-\varphi_i)(w_i - a^W) - \varphi_i c_W)}{2((1-\varphi_i)w_i + b)}. \quad (38)$$

I denote by $(P_i^*; f_i^*; w_i^*)$ the candidate equilibrium solution of (30), (31), and (32). As $f_i + a^C - c_I + k\lambda_i^* = n_i^*\lambda_i^*((1-\varphi_i)(w_i - a^W) - \varphi_i c_W)/2(\alpha S(\lambda_i^*) + (1-\alpha)S)$, by equation (35), replacing for this expression in equation (37), we find that the foreign 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 (33), the deposit fee is given by

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

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 $a_1 \leq 0$, $a_2 \leq 0$, $a_1a_2 - b^2 \geq 0$, $a_1a_3 - c^2 \geq 0$, $a_3a_2 - d^2 \geq 0$ 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 (30), 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 (34), $\left.\frac{\partial A}{\partial f_i}\right|_{p^*} = \left.\frac{\partial C}{\partial f_i}\right|_{p^*}$. Hence,

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

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

$$b = -\frac{1}{t}\left.\frac{\partial C}{\partial f_i}\right|_{p^*} \leq 0.$$

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 (31), 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 (25), I obtain that $\frac{\partial^2 \gamma_i}{\partial^2 f_i} = \frac{\alpha}{2t}\frac{\partial \lambda_i^*}{\partial f_i}\frac{F(\lambda_i^*)}{\lambda_i^*}$. As

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

I can compute the second derivative of A with respect to f_i at p^*. This yields[55]

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

Hence,

$$a_2 = \left.\frac{\partial^2 \pi_i}{\partial^2 f_i}\right|_{p^*} = \frac{-1}{t}\left(\left.\frac{\partial C_i}{\partial f_i}\right|_{p^*}\right)^2 - \alpha\gamma_i\left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\frac{F(\lambda_i^*)}{\lambda_i^*}\left(1 + k\varphi_i\left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right).$$

[55] 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 $\partial\lambda_i^*/\partial f_i \geq 0$, 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

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

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 (33), after some simplifications,

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

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 (32), I obtain,

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

As

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

I have that

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

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

$$a_3 = \left.\frac{\partial^2 \pi_i}{\partial^2 w_i}\right|_{p^*} = (1-\varphi_i)\gamma_i \left.\frac{\partial n_i^*}{\partial w_i}\right|_{p^*} - \frac{1}{t}\left(\left.\frac{\partial C_i}{\partial w_i}\right|_{p^*}\right)^2 - \gamma_i \frac{\alpha F(\lambda_i^*)}{\lambda_i^*}k\varphi_i \left(\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\right)^2.$$

As $\partial n_i^*/\partial w_i \leq 0$ and $\partial \lambda_i^*/\partial w_i \leq 0$, it follows that $a_3 \leq 0$.

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

$$a_1 a_2 - b^2 = \frac{\alpha\gamma_i F(\lambda_i^*)}{t\lambda_i^*} \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*} \left(1 + \varphi_i k \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right) \geq 0,$$

$$a_1 a_3 - c^2 = \frac{-(1-\varphi_i)\gamma_i}{t} \left(\left.\frac{\partial n_i^*}{\partial w_i}\right|_{p^*}\right) + \gamma_i \frac{F(\lambda_i^*)}{t\lambda_i^*} k\alpha\varphi_i \left(\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\right)^2 \geq 0,$$

$$\begin{aligned}
a_3 a_2 - d^2 \quad = \quad & -\frac{1}{t}\left(\left.\frac{\partial C_i}{\partial f_i}\right|_{p^*}\right)^2 (1-\varphi_i)\gamma_i \left.\frac{\partial n_i^*}{\partial w_i}\right|_{p^*} + \frac{\alpha\gamma_i}{t}\frac{F(\lambda_i^*)}{\lambda_i^*} k\varphi_i \left(\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\right)^2 \left(\left.\frac{\partial C_i}{\partial f_i}\right|_{p^*}\right)^2 \\
& -\alpha\left(\gamma_i\right)^2 \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*} \frac{F(\lambda_i^*)}{\lambda_i^*} \left(1 + \varphi_i k \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right) (1-\varphi_i) \left.\frac{\partial n_i^*}{\partial w_i}\right|_{p^*} \\
& +\frac{\gamma_i}{t}\left(\left.\frac{\partial C_i}{\partial w_i}\right|_{p^*}\right)^2 \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*} \frac{\alpha F(\lambda_i^*)}{\lambda_i^*} \left(1 + \varphi_i k \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right) \\
& -\frac{2\gamma_i}{t} \left.\frac{\partial C_i}{\partial f_i}\right|_{p^*} \left.\frac{\partial C_i}{\partial w_i}\right|_{p^*} \left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*} \frac{\alpha F(\lambda_i^*)}{\lambda_i^*} \left(1 + \varphi_i k \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right) \\
& -\left(\gamma_i\right)^2 \left(\left.\frac{\partial \lambda_i^*}{\partial w_i}\right|_{p^*}\right)^2 \left(\frac{\alpha F(\lambda_i^*)}{\lambda_i^*}\right)^2 \left(1 + \varphi_i k \left.\frac{\partial \lambda_i^*}{\partial f_i}\right|_{p^*}\right).
\end{aligned}$$

All the terms except the last one are positive. We have to assume that the first terms are high enough, such that $a_3 a_2 - d^2 \geq 0$. Finally, from the rule of Sarrus,

$$\det H = a_1 a_2 a_3 + 2bdc - c^2 a_2 - b^2 a_3 - d^2 a_1,$$

that is,

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

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

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

I will assume that the distribution of transaction prices is chosen such that this condition is verified.

Appendix C: Proof of Lemma 4 and 5.

Appendix C-1: proof of Lemma 4. I give the conditions that must be verified by the interchange fees such that consumers use both cash and cards at the equilibrium.

First, let us determine the conditions under which consumers do not use cash at the equilibrium when merchants accept cards. From Proposition 1, if $f_i^* \leq \delta l \sqrt{r((1-\varphi_i)w_i^* + b)/2}$, the consumer pays all his transactions by debit card when he meets a merchant who accepts them. 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

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

Hence,

$$\frac{\partial f_i^*}{\partial a^C} = \frac{-1 - k\varphi_i \left[(\partial \lambda_i^*/\partial m)(\partial m/\partial a^C) + (\partial \lambda_i^*/\partial \alpha)(\partial \alpha/\partial a^C) \right]}{1 + k\varphi_i (\partial \lambda_i^*/\partial f_i)}. \tag{39}$$

As $\partial \lambda_i^*/\partial f_i \geq 0$, $\partial \lambda_i^*/\partial \alpha \leq 0$, $\partial \alpha/\partial a^C \leq 0$, $\partial m/\partial a^C = -1$, 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 $\partial\ f_i^*/\partial a^C \leq 0$.[56] Hence, there exists a level of interchange fee $\widehat{a_i^C}(a^W)$ such that $f_i^* > \delta l \sqrt{r((1-\varphi_i)w_i^* + b)/2}$ if $a^C < \widehat{a_i^C}(a^W)$ and $f_i^* \leq \delta l \sqrt{r((1-\varphi_i)w_i^* + b)/2}$ otherwise. As a consequence, using the result of Proposition 1, we conclude that the consumers of bank i do not use cash if $a^C \geq \widehat{a_i^C}(a^W)$ when they meet a merchant who accepts debit cards.

Second, let us determine the conditions under which consumers use only cash at the equilibrium. Assume that $a^W \geq c_W$. Then $w_i^* \geq c_W$. Because of assumption (A3), we have

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

As $f_i^* + a^C + k\varphi_i \lambda_i^* = c_I$ and $-\beta c_I \leq -\beta a^C$, we have that $f_i^* + a^C + k\varphi_i \lambda_i^* < \overline{\lambda}(v_i - \beta(a^C + c_A) + \sqrt{r((1-\varphi_i)w_i^* + b)/2S})$. As $a^C \geq 0$ and $k\varphi_i \lambda_i^* \geq 0$, it follows that $f_i^* < \overline{\lambda}(v_i - \beta(a^C + c_A) + \sqrt{r((1-\varphi_i)w_i^* + b)/2S})$. Therefore, (Condition A-1) is verified. 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, consumers use a combination of cash and card payments at the equilibrium if and only if $a^C < \widehat{a_i^C}(a^W)$.

[56] As $m = a^C + c_A$, we have that $v_i - \beta m = v_i - \beta(a^C + c_A)$. The surcharge rate is sufficiently small in equilibrium if $v_i - \beta m > 0$, that is if $v_i > \beta(a^C + c_A)$. As $a^C \leq c_I$, we have $\beta(a^C + c_A) \leq \beta(c_I + c_A)$. By assumption (A5), we have that $v_i > \beta(c_I + c_A)$, which implies that the surcharge rate is sufficiently small at the equilibrium.

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 card fee, the merchant fee and the share of merchants who accept cards, that is

$$\frac{d\lambda_i^*}{da^C} = \underbrace{\frac{\partial \lambda_i}{\partial f_i}\frac{\partial f_i^*}{\partial a^C}}_{A} + \underbrace{\frac{\partial \lambda_i^*}{\partial m}\frac{\partial m}{\partial a^C}}_{B} + \underbrace{\frac{\partial \lambda_i^*}{\partial \alpha}\frac{\partial \alpha}{\partial a^C}}_{C}. \tag{40}$$

As $\partial\lambda_i^*/\partial f_i \geq 0$ and $\partial f_i/\partial a^C \leq 0$ if β is sufficiently small, term A in (40) is negative. As $\partial\lambda_i^*/\partial m \geq 0$, $\partial m/\partial a^C = 1$, $\partial\lambda_i^*/\partial\alpha \leq 0$ and $\partial\alpha/\partial a^C \leq 0$, terms B and C in (40) are positive. Term B is very small as I assumed that the effect of the surcharge on the behavior of the consumer is negligible compared to other effects. Term C depends on the sensitivity of the share of merchants who accept cards to the interchange fee. If the sensitivity of the share of merchants who accept cards to the interchange fee is small compared to the sensitivity of the consumer fee to the interchange, then $d\lambda_i^*/da^C \leq 0$. 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 in this case, 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 if the sensitivity of α^* to the interchange fee is small, and they 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

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

and

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

As $m = a^C + c_A$ and $w_i = (\varphi_i/(1-\varphi_i))c_W + a^W$, using (25), (27), and (24), the total derivatives of the costs with respect to the interchange fees are

$$\frac{dC_i^*}{da^C} = \underbrace{\frac{\partial f_i^*}{\partial a^C}\int\limits_{\lambda_i(a^C,a^W)}^{\overline{\lambda}}\frac{F(T)}{T}dT}_{A} + \underbrace{\beta\alpha^*\int\limits_{\lambda_i(a^C,a^W)}^{\overline{\lambda}}F(T)dT}_{B} + \underbrace{\frac{\partial C_i^*}{\partial \alpha}\frac{\partial \alpha^*}{\partial a^C}}_{C}, \tag{41}$$

and

$$\frac{dC_i^*}{da^W} = (1-\varphi_i)n_i^* - k\alpha^*\varphi_i\frac{\partial \lambda^*}{\partial a^W}\int\limits_{\lambda^*}^{\overline{\lambda}}\frac{F(T)}{T}dT. \tag{42}$$

We have that $\partial C_i^*/\partial a^W \geq 0$, as $(1-\varphi_i)n_i^* \geq 0$ and $\partial\lambda^*/\partial a^W \leq 0$.

The sign of $\partial C_i^*/\partial a^C$ depends on the sign of terms A, B, and C. As $\partial f_i^*/\partial a^C \leq 0$, term A is negative. As by assumption β is small, term B is positive but small. Term C is positive if the card fee is small, as in this case $\partial C_i^*/\partial\alpha \leq 0$ and $\partial\alpha^*/\partial a^C \leq 0$. It follows that $\partial C_i^*/\partial a^C \leq 0$ if the sensitivity of α to a^C is small enough, as $|A|$ is higher than C. Otherwise, $\partial C_i^*/\partial a^C \geq 0$.

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), the derivative of the merchant surplus with respect to the interchange fee on withdrawals is given by

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

Notice that, as $c_M^* = (1-\beta)(a^C + c_A)$, we have that $E(c_M/c_M \geq c_M^*) \geq (1-\beta)(a^C + c_A)$. As $\partial\lambda^*/\partial a^W \leq 0$ from Lemma 5, it follows that the merchant surplus is increasing with the interchange fee on withdrawals.

From (10), the derivative of the merchant surplus with respect to the interchange fee on card payments is given by

$$\frac{dMS}{da^C} = -\alpha^*(1-\beta)\int_{\lambda^*}^{\overline{\lambda}} F(T)dT - \alpha^* \frac{\partial \lambda^*}{\partial a^C}\left[E(c_M/c_M \geq c_M^*) - (1-\beta)(a^C + c_A)\right] F(\lambda^*)$$
$$-(1-\alpha^*)S\frac{dE(c_M/c_M \leq c_M^*)}{da^C} - \alpha^* S(\lambda^*)\frac{dE(c_M/c_M \geq c_M^*)}{da^C} + \frac{d\alpha^*}{da^C}\frac{\partial MS}{\partial \alpha},$$

where

$$\frac{\partial MS}{\partial \alpha} = S\left[E(c_M/c_M \leq c_M^*) - (1-\beta)(a^C + c_A)\right] + S(\lambda^*)\left[(1-\beta)(a^C + c_A) - E(c_M/c_M \geq c_M^*)\right].$$

As $c_M^* = (1-\beta)(a^C + c_A)$, we have $E(c_M/c_M \leq c_M^*) \leq (1-\beta)(a^C + c_A)$ and $(1-\beta)(a^C + c_A) \leq E(c_M/c_M \geq c_M^*)$. It follows that $\partial MS/\partial \alpha \leq 0$. As $dE(c_M/c_M \leq c_M^*)/da^C = -dE(c_M/c_M \geq c_M^*)/da^C$, we have

$$\frac{dMS}{da^C} = \underbrace{-\alpha^*(1-\beta)\int_{\lambda^*}^{\overline{\lambda}} F(T)dT}_{\text{Term A}} - \underbrace{\alpha^* F(\lambda^*)\frac{\partial \lambda^*}{\partial a^C}(E(c_M/c_M \geq c_M^*) - (1-\beta)(a^C + c_A))}_{\text{Term B}}$$
$$+\underbrace{\frac{dE(c_M/c_M \leq c_M^*)}{da^C}\left[-S + \alpha^*(S - S(\lambda^*))\right]}_{\text{Term C}} + \underbrace{\frac{d\alpha^*}{da^C}\frac{\partial MS}{\partial \alpha}}_{\text{Term D}}.$$

Term A is negative. As $\partial \lambda^*/\partial a^C \leq 0$ if the sensitivity of the share of merchants is small, term B is positive. We also have $dE(c_M/c_M \leq c_M^*)/da^C \geq 0$.As $dE(c_M/c_M \leq c_M^*)/da^C \geq 0$, term C is negative. As $\partial MS/\partial \alpha \leq 0$ and $\partial \alpha^*/\partial a^C \geq 0$, term D is negative.

The impact of the interchange fee on card payments on merchant surplus depends on two effects. First, a rise in the interchange fee on card payments increases the cost of accepting debit cards. This raises the costs borne by merchants who accept debit cards when consumers use them (term A). This effect also decreases the share of merchants who accept cards, which increases merchant surplus (term D), but increases the costs of cash borne by merchants (term C). Second, a rise in the interchange fee on card payments increases the value of debit card payments, which reduces the costs of the merchants who accept cards (term B), because, for these merchants, debit cards are less costly to accept than cash.

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

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

where

$$\frac{\partial P^*}{\partial \alpha} = k(1-2\varphi)\int_{\lambda^*}^{\overline{\lambda}} F(T)dT + k\varphi\lambda^* \int_{\lambda^*}^{\overline{\lambda}} \frac{F(T)}{T}dT \geq 0.$$

Notice that if the costs of cash for banks are high, $\partial P^*/\partial \alpha$ is high.

From (41),

$$\frac{dC^*}{da^C} = \alpha^*(-1 - k\varphi\frac{\partial \lambda^*}{\partial a^C})\int_{\lambda^*}^{\overline{\lambda}} \frac{F(T)}{T}dT + \beta\alpha^* \int_{\lambda^*}^{\overline{\lambda}} F(T)dT + \frac{\partial C^*}{\partial \alpha}\frac{d\alpha^*}{da^C}.$$

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

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

and from (42),

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

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

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

and

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

An increase in the interchange fee on withdrawals has a negative impact on consumer surplus, as it increases his transaction costs (the card fee and the withdrawal fee), which is only compensated by a reduction in the demand for foreign withdrawals. The impact of an increase in the interchange fee on card payments is more complex as it reduces the card fee, which benefits consumers, but it reduces the share of merchants who accept cards, which may be detrimental to consumers.

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

We start by studying the case of the interchange fee on withdrawals. We have that

$$\left.\frac{dTUS}{da^W}\right|_{IF} = -(1-\varphi)n^* - \alpha^*(E(c_M/c_M \geq c_M^*) - (1-\beta)((a^C)^P + c_A)) \left.\frac{\partial\lambda^*}{\partial a^W}\right|_{IF} F(\lambda^*).$$

The first term of the previous equation is negative and represents the volume of foreign withdrawals. The second term is positive and represents the costs savings that are made by merchants when consumers substitute debit cards for cash because of an increase in a^W. It follows that $\left.\frac{dTUS}{da^W}\right|_{IF} \leq 0$ if the volume of foreign withdrawals is high at the profit maximising interchange fee, and if the share of merchants who accept cards is small. As by assumption W is concave in a^W, we conclude that, for a given $(a^C)^\pi$, the interchange fee on withdrawals is too high to maximise the total user surplus if the volume of foreign withdrawals is high, and if the share of merchants who accept cards is small.

Now let us study the case of the interchange fee on card payments. We have that

$$\begin{aligned}
\left.\frac{dTUS}{da^C}\right|_{IF} &= -\frac{d\alpha^*}{da^C}\left[k\varphi(\alpha^* S(\lambda^*) + (1-\alpha^*)S) + k\varphi\lambda^* \int_{\lambda^*}^{\bar{\lambda}} \frac{F(T)}{T} dT\right] \\
&\quad - \frac{\partial C^*}{\partial \alpha}\frac{d\alpha^*}{da^C} + \alpha^* \int_{\lambda^*}^{\bar{\lambda}} \frac{F(T)}{T} dT - \alpha^* \int_{\lambda^*}^{\bar{\lambda}} F(T) dT \\
&\quad + \frac{d\alpha^*}{da^C}\frac{\partial MS}{\partial \alpha} + \frac{dE(c_M/c_M \leq c_M^*)}{da^C}\left[-S + \alpha^*(S - S(\lambda^*))\right] \\
-\alpha^* F(\lambda^*)\frac{\partial \lambda^*}{\partial a^C}(E(c_M/c_M &\geq c_M^*) - (1-\beta)(a^C + c_A)).
\end{aligned}$$

All the terms of $\left.\frac{dTUS}{da^C}\right|_{IF}$ are positive, except $-\frac{\partial C^*}{\partial \alpha}\frac{d\alpha^*}{da^C}$, which may be negative, and $-\alpha^* \int_{\lambda^*}^{\bar{\lambda}} F(T)dT$. The term $-\frac{\partial C^*}{\partial \alpha}\frac{d\alpha^*}{da^C}$ represents the reduction in consumer transaction costs that is due to lower acceptance of payment cards. The term $-\alpha^* \int_{\lambda^*}^{\bar{\lambda}} F(T)dT$ represents the loss borne by merchants when the merchant fee increases. If the costs of cash are small for merchants, the share of merchants who accept cards is small. As a result, this term is small at the profit maximising interchange fees. If $\frac{d\alpha^*}{da^C}$ is small and if k (the costs of cash for banks) is high, we have $\left.\frac{dTUS}{da^C}\right|_{IF} \geq 0$. It follows that $\left.\frac{dW}{da^C}\right|_{IF} \geq 0$. Hence, for a given level of $(a^W)^\pi$, as W is concave in a^C by assumption, social welfare can be increased by raising the interchange fee on card payments.

Otherwise, if the costs of cash are high for merchants, this term is negative and can be significantly high. If $\frac{d\alpha^*}{da^C}$ and $\frac{d\lambda^*}{da^C}$ is small, $\left.\frac{dW}{da^C}\right|_{IF} \leq 0$, and social welfare can be increased by lowering the level of interchange fees.

The results are summarized in the following table (SW standing for social welfare)

<table>
<tr><th rowspan="2">Costs of cash for merchants</th><th colspan="2">Costs of cash for banks</th></tr>
<tr><th>Low</th><th>High</th></tr>
<tr><td rowspan="2">Low</td><td>SW can be increased by lowering (resp. raising) the IF on card payment if $dC^*/d\alpha$ is higher (resp. lower) than $dMS/d\alpha$ at the profit maximising IF.</td><td>SW can be increased by raising the IF on card payments.</td></tr>
<tr><td colspan="2">SW can be increased by lowering the interchange fee on withdrawals if the number of foreign withdrawals is relatively high.</td></tr>
<tr><td rowspan="2">High</td><td>SW can be increased by lowering the IF on card payments if the sensitivity to the IF of the share of merchants who accept cards is small. and if $d\lambda^*/da^C$ is small.</td><td>No clear policy recommendation.</td></tr>
<tr><td colspan="2">SW can be increased by increasing the interchange fee on withdrawals if the number of foreign withdrawals is not too high at the profit maximising interchange fees.</td></tr>
</table>

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: Asymmetries between issuers. Assume that bank 1 owns all the ATM network. 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

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

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

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

The profit maximising interchange fee on withdrawals verifies

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

Appendix F-2: 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-3: ATM Deployment decisions Taking the derivative of (20) with respect to ρ_i, I obtain that

$$\begin{aligned}
& \frac{\partial \lambda_i^*}{\partial \rho_i}\sqrt{\frac{r((1-\varphi_i)w_i+b)}{2(\alpha S(\lambda_i^*)+(1-\alpha)S)}}\left(1-\frac{\alpha F(\lambda_i^*)\lambda_i^*}{2(\alpha S(\lambda_i^*)+(1-\alpha)S)}\right) \\
= \; & \frac{-\lambda_i^*}{2}\sqrt{\frac{r}{(\alpha S(\lambda_i^*)+(1-\alpha)S)((1-\varphi_i)w_i+b)}}\left(\frac{-\rho_j w_i}{\rho_1+\rho_2}+b'(\rho_1+\rho_2)\right).
\end{aligned}$$

As by assumption b' is negative, the expression in the right side of the equality is positive. It follows that the threshold above which consumers pay by debit card increases with the number of ATMs deployed by his bank. A similar proof shows that n_i^* increases with the number of ATMs deployed by the consumer's bank.

From (1) and from the envelop's theorem, we have that

$$\left.\frac{\partial C_i^*(\lambda_i^*)}{\partial \rho_i}\right|_{(\lambda_i^*,n_i^*)} = \frac{1}{2}\sqrt{\frac{2r(\alpha S(\lambda_i^*)+(1-\alpha)S)}{((1-\varphi_i)w_i+b)}}\left(\frac{-\rho_j w_i}{\rho_1+\rho_2}+b'(\rho_1+\rho_2)\right).$$

As by assumption b' is negative, the consumer's transaction costs decrease with the number of ATMs deployed by his bank.

If we consider ATM deployment costs, bank i's profit is given by

$$\pi_i = 2t(\gamma_i^*)^2 + n_j^*(\frac{\rho_i}{\rho_1+\rho_2})(a^W - c_W) - k(\frac{\rho_i}{\rho_1+\rho_2})\left[\alpha^* S(\lambda_j^*)+(1-\alpha^*)S\right] - DC(\rho_i).$$

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

$$\begin{aligned}
& 4t\frac{\partial \gamma_i^*}{\partial \rho_i}\gamma_i^* + \frac{\partial n_j^*}{\partial \rho_i}(\frac{\rho_i}{\rho_1+\rho_2})(a^W-c_W) + n_j^*\frac{\rho_j}{(\rho_1+\rho_2)^2}(a^W-c_W) \\
= \; & \alpha^* k(\frac{\rho_i}{\rho_1+\rho_2})\frac{\partial \lambda_j^*}{\partial \rho_i}F(\lambda_j^*) + k\frac{\rho_j}{(\rho_1+\rho_2)^2}\left[\alpha^* S(\lambda_j^*)+(1-\alpha^*)S\right] + DC'(\rho_i).
\end{aligned}$$

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:

$$\begin{aligned}
& \frac{1}{2}\frac{\partial n^*}{\partial \rho}(a^W-c_W) + \frac{n^*}{4\rho}(a^W-c_W) \\
= \; & \frac{k\alpha^*}{2}\frac{\partial \lambda^*}{\partial \rho}F(\lambda^*) + \frac{k}{4\rho}\left[\alpha^* S(\lambda^*)+(1-\alpha^*)S\right] + DC'(\rho).
\end{aligned}$$