

Rochet, Jean-Charles; Wright, Julian

**Working Paper**

## Credit card interchange fees

ECB Working Paper, No. 1138

**Provided in Cooperation with:**

European Central Bank (ECB)

*Suggested Citation:* Rochet, Jean-Charles; Wright, Julian (2009) : Credit card interchange fees, ECB Working Paper, No. 1138, European Central Bank (ECB), Frankfurt a. M.

This Version is available at:

<https://hdl.handle.net/10419/153572>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



EUROPEAN CENTRAL BANK

EUROSYSTEM

---

**RETAIL PAYMENTS:  
INTEGRATION AND INNOVATION**

---

**WORKING PAPER SERIES**

**NO 1138 / DECEMBER 2009**

**CREDIT CARD  
INTERCHANGE FEES**

Jean-Charles Rochet  
and Julian Wright



EUROPEAN CENTRAL BANK

EUROSYSTEM



## WORKING PAPER SERIES

NO 1138 / DECEMBER 2009

---

**RETAIL PAYMENTS:  
INTEGRATION AND INNOVATION**

---

### CREDIT CARD INTERCHANGE FEES<sup>1</sup>

Jean-Charles Rochet<sup>2</sup> and

Julian Wright<sup>3</sup>



In 2009 all ECB publications feature a motif taken from the €200 banknote.

This paper can be downloaded without charge from  
<http://www.ecb.europa.eu> or from the Social Science Research Network  
electronic library at [http://ssrn.com/abstract\\_id=1522025](http://ssrn.com/abstract_id=1522025).



<sup>1</sup> We benefited from the helpful comments of conference participants at the AEA January 2009 meeting in San Francisco (especially our discussant, Charlie Kahn) and two Conferences on Payments Economics at the Central Bank of Norway (November 2008) and the Central Bank of Canada (November 2008).

<sup>2</sup> Toulouse School of Economics, Manufacture de Tabacs, 21 allées de Brienne - 31000 Toulouse, France; e-mail: [rochet@cict.fr](mailto:rochet@cict.fr)

<sup>3</sup> Department of Economics, National University of Singapore, 21 Lower Kent Ridge Road, Singapore 119077; e-mail: [jwright@nus.edu.sg](mailto:jwright@nus.edu.sg)

## Retail payments: integration and innovation

“Retail payments: integration and innovation” was the title of the joint conference organised by the European Central Bank (ECB) and De Nederlandsche Bank (DNB) in Frankfurt am Main on 25 and 26 May 2009. Around 200 high-level policy-makers, academics, experts and central bankers from more than 30 countries of all five continents attended the conference, reflecting the high level of interest in retail payments.

The aim of the conference was to better understand current developments in retail payment markets and to identify possible future trends, by bringing together policy conduct, research activities and market practice. The conference was organised around two major topics: first, the economic and regulatory implications of a more integrated retail payments market and, second, the strands of innovation and modernisation in the retail payments business. To make innovations successful, expectations and requirements of retail payment users have to be taken seriously. The conference has shown that these expectations and requirements are strongly influenced by the growing demand for alternative banking solutions, the increasing international mobility of individuals and companies, a loss of trust in the banking industry and major social trends such as the ageing population in developed countries. There are signs that customers see a need for more innovative payment solutions. Overall, the conference led to valuable findings which will further stimulate our efforts to foster the economic underpinnings of innovation and integration in retail banking and payments.

We would like to take this opportunity to thank all participants in the conference. In particular, we would like to acknowledge the valuable contributions of all presenters, discussants, session chairs and panellists, whose names can be found in the enclosed conference programme. Their main statements are summarised in the ECB-DNB official conference summary. Twelve papers related to the conference have been accepted for publication in this special series of the ECB Working Papers Series.

Behind the scenes, a number of colleagues from the ECB and DNB contributed to both the organisation of the conference and the preparation of this conference report. In alphabetical order, many thanks to Alexander Al-Haschimi, Wilko Bolt, Hans Brits, Maria Foskolou, Susan Germain de Urday, Philipp Hartmann, Päivi Heikkinen, Monika Hempel, Cornelia Holthausen, Nicole Jonker, Anneke Kosse, Thomas Lammer, Johannes Lindner, Tobias Linzert, Daniela Russo, Wiebe Ruttenberg, Heiko Schmiedel, Francisco Tur Hartmann, Liisa Väisänen, and Pirjo Väkeväinen.

Gertrude Tumpel-Gugerell  
Member of the Executive Board  
European Central Bank

Lex Hoogduin  
Member of the Executive Board  
De Nederlandsche Bank

© European Central Bank, 2009

### Address

Kaiserstrasse 29  
60311 Frankfurt am Main, Germany

### Postal address

Postfach 16 03 19  
60066 Frankfurt am Main, Germany

### Telephone

+49 69 1344 0

### Website

<http://www.ecb.europa.eu>

### Fax

+49 69 1344 6000

*All rights reserved.*

*Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the author(s).*

*The views expressed in this paper do not necessarily reflect those of the European Central Bank.*

*The statement of purpose for the ECB Working Paper Series is available from the ECB website, <http://www.ecb.europa.eu/pub/scientific/wps/date/html/index.en.html>*

ISSN 1725-2806 (online)

# CONTENTS

Abstract	4
1 Introduction	4
2 The model	6
3 Analysis and policy implications	9
3.1 When do retailers accept credit cards?	10
3.2 Analysis of retailer prices and consumer surplus	12
3.3 Policy implications	14
4 Alternative specifications	16
4.1 Allowing store credit for ordinary purchases	17
4.2 Endogenizing card membership decisions	20
5 Conclusions	23
Appendix	26
References	27
European Central Bank Working Paper Series	28



## Abstract

We build a model of credit card pricing that explicitly takes into account credit functionality. We show that a monopoly card network always selects an interchange fee that exceeds the level that maximizes consumer surplus. If regulators only care about consumer surplus, a conservative regulatory approach is to cap interchange fees based on retailers' net avoided costs from not having to provide credit themselves. In the model, this always raises consumer surplus compared to the unregulated outcome, sometimes to the point of maximizing consumer surplus.

## 1 Introduction

Even though payment cards are gradually becoming the most popular and most efficient means of payments in many countries, there is a growing suspicion surrounding the pricing of credit cards. Retailers complain that the fees they have to pay to accept credit card transactions are out of proportion with the costs incurred by banks. Some competition authorities and central banks have suggested banks provide consumers with exaggerated incentives to use their credit cards, to the detriment of other means of payments like cash and debit cards which they believe to be more efficient. The usual suspects are the interchange fees, the transfer fees paid by the bank of the retailer to the banks of the cardholders, which are often considerably higher than those for **debit cards**.<sup>1</sup> In the past several years, there have been more than 50 lawsuits concerning interchange

---

<sup>1</sup>Unregulated credit card interchange fees are typically between 1% to 2% of transaction value, whereas debit card interchange fees are typically between 0% and 1%. See, for instance, Charts 2 and 3 in Weiner and Wright (2005).

fees filed by merchants and merchant associations against card networks in the United States, while in about 20 countries public authorities have taken regulatory actions related to interchange fees and investigations are proceeding in many more (Bradford and Hayashi, 2008).

Given the obvious importance of understanding how interchange fees should be set, this article analyzes credit card interchange fee determination to see whether there are grounds for regulatory intervention, and if so, in what form. The point of departure from the existing literature is to model *credit cards* explicitly. An existing literature models price determination in payment cards networks, initiated by Schmalensee (2002), Rochet and Tirole (2002) and Wright (2003).<sup>2</sup> The models in this literature have essentially focused on the choice between payment cards (which could just as well be debit cards) and cash (or checks). We contribute to this literature by extending the models to allow a separate role for the credit functionality of credit cards, thereby allowing us to discuss credit card interchange fees specifically.<sup>3</sup>

In our model, credit cards can be used for two types of transactions — “ordinary purchases” for regular convenience usage for which cash (or a debit card) are assumed to provide identical benefits, and for “credit purchases” where credit is necessary for purchases to be realized. Credit purchases include a range of different types of purchases (such as unplanned purchases, impulse purchases and large purchases) for which the consumer does not have the cash or funds immediately available to complete the purchase **or for purchases for which the deferment of payment facilitates the transaction.** For ordinary purchases, we assume credit cards are inefficient given we assume there are additional costs of transacting with credit cards. As a result, card networks which maximize profit by maximizing the number of card transactions have an incentive to encourage over-usage of credit cards by convenience users (even when these consumers do not need the credit facility) provided merchants still accept such credit card transactions. A card network does this by setting interchange fees high enough to induce issuers to offer rewards and cash back bonuses (equivalent to negative fees). On the other hand, the alternative to using credit cards for credit purchases is the direct provision of credit by merchants or “store credit”,

---

<sup>2</sup>See also Baxter (1983) for a much earlier treatment, and Rochet (2003) for a survey of the literature.

<sup>3</sup>Among the few papers to model explicitly the credit functionality are Chakravorty and To (2007) and more recently Bolt and Chakravorty (2008), although these papers do not focus on the determinant of interchange fees for credit cards, nor the regulation of these fees.

which is assumed to be relatively inefficient. Since consumers do not internalize retailers' cost savings from avoiding direct provision of credit and since merchants cannot distinguish the type of consumer they face, there is also a case for setting a relatively high interchange fee so that consumers **that wish to rely on** credit are induced to use credit cards when it is efficient for them to do so. For this reason, to maximize consumer surplus (including the surplus of cash customers) may require setting an interchange fee which **induces excessive usage of credit cards for ordinary purchases**.

Taking into account both types of transactions, a monopoly card network always sets its interchange fee too high in our setting. Thus, if regulators only care about (short-run) consumer surplus, our theory can provide a rationalization for placing a cap on interchange fees.<sup>4</sup> The theory suggests one of two possible caps will maximize consumer surplus. Depending on the relative costs and benefits of the different instruments, the cap should either be based on the issuers' costs (to avoid excessive usage of cards for **ordinary purchases**) or on merchants' net avoided costs from not having to provide credit directly (so that consumers use their cards efficiently for credit purchases). Since evaluating which of the two options gives higher consumer surplus is informationally very demanding, a conservative regulatory approach would be to cap interchange fees using the maximum of these two levels, which **is likely to be** the latter option. In our model, this always raises consumer surplus compared to the unregulated outcome, and **will sometimes** result in the best outcome for consumers. In contrast, using issuer costs to regulate interchange fees is **realistically only likely to give** a lower bound of possible interchange fees that maximize consumer surplus.

## 2 The Model

There is a continuum of consumers (of total mass normalized to one) with quasi-linear preferences. They spend their income on a composite good taken as a numéraire and on retail goods costing  $\gamma$  to produce. There are two payment technologies: cash (which could also capture cheques and debit cards<sup>5</sup>) and credit cards. Credit cards are assumed to be more costly but allow consumers

<sup>4</sup>In focusing on consumer surplus, we ignore the need for issuers to recover fixed costs and the effect this has on entry incentives (and therefore, on long-run consumers surplus). We also ignore the need to get consumers to internalize the effect of their decisions on the profit of issuers so as to maximize total welfare. As Rochet and Tirole (2008) show, taking these effects into account justifies higher interchange fees.

<sup>5</sup>Since the focus of this paper is on the provision of consumer credit, we do not introduce any differences between **pure payment** technologies. Such differences are discussed at length in the literature mentioned in the introduction, e.g. Schmalensee (2002), Rochet and Tirole (2002) and Wright (2003).



to purchase on credit. They are held by a fraction  $x$  of consumers, where  $x$  is initially taken as given.<sup>6</sup> Consumers purchase one unit of the retail good (what we call “ordinary purchases”) providing them with utility  $u_0$  with  $u_0 > \gamma$ . In addition, with probability  $\theta$ , they also receive utility  $u_1$  from consuming another unit of the retail good (what we call “credit purchases”). We assume that merchants cannot bundle the two transactions and also cannot distinguish between “ordinary” and “credit” purchases. We assume that each consumer has always enough cash to pay for his ordinary purchases, but must rely on credit for credit purchases.

Each retailer can directly provide credit to the consumer (“store credit”), but this entails a cost  $c_B$  for the consumer (buyer) and cost  $c_S$  to the retailer (seller). The cost  $c_S$  is the same for all credit purchases from a given seller, but  $c_B$  is transaction specific, and is observed by the consumer only when he is in the store.  $c_B$  is drawn from a continuous distribution with the cumulative distribution function  $H$ . We assume the distribution has full support over some range  $(\underline{c}_B, \bar{c}_B)$  where  $\underline{c}_B$  is sufficiently negative, **such that cardholders will sometimes choose to use store credit**, and  $\bar{c}_B$  is positive but not too high (**in comparison with  $u_1 - \gamma$** ), such that consumers will always prefer to make the credit purchase (**even if they have to pay with store credit**) than not buy at all.<sup>7</sup> **Initially, we assume  $c_B$  only arises with respect to credit purchases, and represents the net cost of using store credit rather than credit cards for credit purchases, which is why we allow the possibility of negative draws of  $c_B$ . In section 4 we will analyze what happens when  $c_B$  is instead interpreted as a transaction cost that arises equally for ordinary purchases and credit purchases.**

Given the costs of store credit, accepting credit cards is a potential means for merchants to reduce their transaction costs of accepting credit purchases and to increase the quality of service to buyers. The cost of a credit transaction is  $c_A$  for the bank of the merchant (which is called the acquirer of the transaction) and  $c_I$  for the bank of the cardholder (the issuer of the card). The total cost of a credit card transaction is thus  $c = c_A + c_I$ . Bank fees for credit card transactions are denoted  $f$  for consumers and  $m$  for merchants. When  $f < 0$  (cash back bonuses) consumers prefer to also use their credit cards for ordinary purchases, which given our assumptions is socially wasteful.<sup>8</sup> For each credit card transaction, an interchange fee  $a$  is paid

<sup>6</sup>In section 4 we extend the model to allow consumers to decide whether to join the card network or not.

<sup>7</sup>**Without this assumption, credit card acceptance would not just add value by reducing the transaction costs of arranging credit; it would also lead to an increase in the volume of credit purchases, which would provide an additional factor affecting efficient interchange fees.**

<sup>8</sup>We assume consumers and **retailers** face no costs of using cash (which can be thought of as a normalization). For ordinary purchases, any costs or benefits **to consumers and retailers** of using cash are assumed to be the same as those obtained from using credit cards (which can be thought of as an approximation).

by the bank of the merchant (the acquirer) to the bank of the consumer (the issuer).

For simplicity, we assume that acquiring merchants is perfectly competitive for banks, which implies that the merchant fee  $m$  is equal to the sum of the acquiring cost  $c_I$  and the interchange fee  $a$ :

$$m = c_I + a. \quad (1)$$

By contrast, we assume that issuers are imperfectly competitive: the cardholder fee  $f$  is equal to the net issuer cost  $c_I - a$  plus a profit margin  $\pi$ , assumed for simplicity to be constant. We thus have:

$$f(a) = c_I - a + \pi. \quad (2)$$

**These assumptions are not made to necessarily capture any intrinsic asymmetry between the nature of competition in issuing and acquiring, the existence of which is an empirical matter. Rather, the main reason for making these particular assumptions on bank competition is that they provide a simple setting in which card networks seek to maximize profit by maximizing the number of card transactions.**

The level of the interchange fee  $a$  has an impact on retailer cost and thus on the retail price<sup>9</sup>  $p(a)$ , which results from competition between retailers. To model competition between retailers, we use the standard Hotelling model: consumers are uniformly distributed on an interval of unit length, with one retailer ( $i = 1, 2$ ) located at each extremity of the interval. Transport cost for consumers is  $t$  per unit of distance. We consider the case where credit card service are provided by a single network (monopoly). Its objective is to maximize total profits of its member banks, which is proportional to total volume of credit card transactions. **As noted above, this simplification is due to our assumption that banks' profit margins are constant (0 for acquirers,  $\pi$  for issuers).**

The timing of our model is as follows:

- The card network sets the interchange fee  $a$  so as to maximize banks' total profit.
- Banks set their fees:  $f(a) = c_I - a + \pi$  for cardholders and  $m(a) = c_A + a$  for retailers.
- Retailers independently choose their card acceptance policies:  $L_i = 1$  if retailer  $i$  accepts credit cards, 0 otherwise.

---

<sup>9</sup>We assume that retailers cannot, or do not want to, charge different retail prices for cash and card payments (i.e. no surcharging).

- After observing  $(L_1, L_2)$ , retailers independently set retail prices  $p_1, p_2$ .
- Consumers observe retail prices and card acceptance policies and select one retailer to patronize.
- Once the consumer is in the store, he buys a first unit of the retail good (“ordinary purchase”), and pays it by cash or a credit card (if he has one).
- Finally, nature decides whether he has an opportunity for a credit purchase (this occurs with probability  $\theta$ ) and in this case, the cost  $c_B$  of using store credit for the buyer is drawn according to the c.d.f.  $H$ , with full support on  $[\underline{c}_B, \bar{c}_B]$ . Cardholders then select their mode of payment.

In the framework above we assumed that consumers sometimes will obtain a negative draw of  $c_B$  (i.e.  $\underline{c}_B < 0$ ). This was one way to ensure store credit is not dominated by credit cards. Despite this, we also assumed store credit is never used for ordinary purchases. We did this by assuming that for ordinary purchases consumers can only choose between cash and credit cards, **consistent with the net cost of using store credit always being positive for ordinary transactions**. This helps to simplify the analysis. **Subsequently, we will show our results continue to hold even when consumers may sometimes prefer store credit for ordinary purchases**. Another reason why store credit may not be dominated by credit cards is that some consumers will not want to hold credit cards unless the costs of doing so are sufficiently subsidized (or they receive sufficient rewards for doing so). We, thus, also analyze an alternative specification in which we endogenize the choice by consumers of whether to hold a credit card in the first place but then assume the cost to consumers of using store credit  $c_B$  **is always positive**. These two extensions are analyzed in Section 4.

### 3 Analysis and policy implications

We first analyze the model to derive optimal interchange fees. This involves determining when retailers will accept credit cards so as to determine what interchange fees a card network will set. We then compare this to the interchange fees that maximize consumer surplus and derive some policy implications.

### 3.1 When Do Retailers Accept Credit Cards?

This section derives the equilibrium behavior of retailers as a function of the fundamental policy variable in our model, namely the interchange fee  $a$ . We first construct a retailer's profit function.

We will use the notation  $L_c = 1$  if  $f < 0$ ,  $L_c = 0$  if  $f \geq 0$  to distinguish whether credit cards are used for ordinary purchases or not. A fraction  $1 - xL_i$  of the time, consumers cannot use credit cards. For ordinary purchases, such consumers must therefore use cash. For credit purchases, which happen with probability  $\theta$ , these consumers will always use store credit. A fraction  $xL_i$  of the time, consumers can also use credit cards. For ordinary purchases, these cardholders will use cash if  $f \geq 0$  (i.e. if  $L_c = 0$ ) and credit cards if  $f < 0$  (i.e. if  $L_c = 1$ ). For credit purchases, which happen with probability  $\theta$ , these cardholders will use store credit if  $c_B < f$  but otherwise will use credit cards. Collecting together a retailer's margins associated with each of these different possibilities, retailer  $i$ 's expected margin per-customer is therefore equal to

$$\begin{aligned} M_i = & (1 - xL_i)(p_i - \gamma + \theta(p_i - \gamma - c_S)) \\ & + xL_i(p_i - \gamma - L_cm + \theta(H(f)(p_i - \gamma - c_S) + (1 - H(f))(p_i - \gamma - m))) \end{aligned}$$

or after simplifying

$$M_i = (1 + \theta)(p_i - \gamma) - \theta c_S - x\Gamma(a)L_i,$$

where

$$\Gamma(a) = L_cm + \theta(m - c_S)(1 - H(f))$$

is the retailer's expected net cost per-cardholder from accepting cards as a function of the interchange fee.

Corresponding to each of the different types of possible transactions, retailer  $i$  offers an expected surplus (ignoring transportation costs) of

$$\begin{aligned} U_i = & (1 - xL_i)(u_0 + \theta u_1 - (1 + \theta)p_i - \theta E(c_B)) \\ & + xL_i \left( u_0 + \theta u_1 - (1 + \theta)p_i - L_cf - \theta \left( \int_{\underline{c}_B}^f c_B dH(c_B) + \int_f^{\bar{c}_B} f dH(c_B) \right) \right) \end{aligned}$$

or after simplifying

$$U_i = u_0 + \theta u_1 - (1 + \theta)p_i - \theta E(c_B) + xS(a)L_i,$$

where

$$S(a) = -L_cf + \theta \int_f^{\bar{c}_B} (c_B - f) dH(c_B)$$

is the expected net consumer surplus per-cardholder from credit card usage as a function of the interchange fee. Determining retailer  $i$ 's market share by finding the indifferent consumer in the normal way, we find<sup>10</sup>

$$s_i = \frac{1}{2} + \frac{(1 + \theta)(p_j - p_i)}{2t} + xS(a) \frac{L_i - L_j}{2t}. \quad (3)$$

Retailer  $i$ 's profit function can then be written  $\pi_i = M_i s_i$ . Let us denote by  $\phi(a)$  the difference between  $S(a)$  and  $\Gamma(a)$ , which we call total user surplus:

$$\phi(a) = S(a) - \Gamma(a) = \theta \int_{c_I + \pi_I - a}^{\bar{c}_B} (c_B + c_S - c - \pi_I) dH(c_B) - (c + \pi_I)L_c. \quad (4)$$

The first term represents total expected surplus of the two users (consumers and retailers) from using credit cards for consumers' credit purchases. The second term is a deadweight loss associated **with the use of credit cards for ordinary purchases**. We are now ready to derive the equilibrium choices of retailers. We first derive equilibrium resulting from price competition between retailers for given card acceptance decisions  $L_1, L_2$ .

**Proposition 1** *For any couple  $L_1, L_2$  of card acceptance decisions, price competition between retailers leads to retail prices such that*

$$(1 + \theta)p_i^* = t + \gamma(1 + \theta) + \theta c_S + x\Gamma(a)L_i + \frac{x}{3}\phi(a)(L_i - L_j). \quad (5)$$

*Retailers' profits are  $\pi_i^* = 2t(s_i^*)^2$ , where retailer  $i$ 's equilibrium market share  $s_i$  is*

$$s_i^* = \frac{1}{2} + \frac{x\phi(a)(L_i - L_j)}{6t}. \quad (6)$$

**Proof of Proposition 1:** See the Appendix.

So as to maximize profit, each retailer chooses to accept cards if it increases its market share, which will be the case whenever  $\phi(a) \geq 0$ . An immediate consequence of Proposition 1 is:

**Proposition 2** *Retailers accept credit cards at equilibrium if and only if expected consumer surplus from card transactions exceeds expected retailer cost. That is,  $L_1^* = L_2^* = 1$  if and only if  $\phi(a) \geq 0$ , i.e.  $S(a) \geq \Gamma(a)$ .*

---

<sup>10</sup>We assume that  $t$  is large enough so that each retailer always has a positive market share at equilibrium ( $0 < s_i < 1$  for all  $i$ ). We also assume that  $u_0 + \theta u_1$  is large enough so that all the market is served ( $U_i > ts_i$  for all  $i$ ).



### 3.2 Analysis of Retailer Prices and Consumer Surplus

This section considers the impact of the interchange fee  $a$  on consumer surplus. Obviously this question only matters in the region where  $S(a) \geq \Gamma(a)$ , i.e. when credit cards are accepted.

Proposition 1 allows us to compute retail prices and consumer surplus as a function of  $a$ . Considering the case where  $L_1 = L_2 = 1$  and thus  $p_1(a) = p_2(a) = p(a)$ , formula (5) gives

$$(1 + \theta)p(a) = t + \gamma(1 + \theta) + \theta c_S + x\Gamma(a). \quad (7)$$

Since  $\Gamma(a)$  (the expected net retailer cost from card transactions) increases in  $a$  we obtain an immediate corollary of Proposition 1.

**Corollary 1** *The equilibrium retail price is an increasing function of the interchange fee  $a$ .*

The expected surplus of cash consumers (those not holding credit cards) is

$$U_{\text{cash}}(a) = u_0 + \theta u_1 - \frac{t}{4} - (1 + \theta)p(a) - \theta E(c_B)$$

and of consumers holding credit cards is

$$U_{\text{credit}}(a) = u_0 + \theta u_1 - \frac{t}{4} - (1 + \theta)p(a) - L_c f - \theta \left( \int_{c_B}^f c_B dH(c_B) + \int_f^{\bar{c}_B} f dH(c_B) \right).$$

If we aggregate the surplus of all consumers (cash consumers and cardholders) and take into account equilibrium prices from (7), we obtain

$$\begin{aligned} CS(a) &= xU_{\text{credit}}(a) + (1 - x)U_{\text{cash}}(a) \\ &= u_0 + \theta u_1 - \gamma(1 + \theta) - \frac{5t}{4} - \theta(E(c_B) + c_S) + x\phi(a). \end{aligned}$$

Thus, aggregate consumer surplus is equal, up to an additive and a positive multiplicative constants, to total user surplus  $\phi(a)$ .

The following three figures show how total user surplus  $\phi$  varies with the level of the interchange fee  $a$ . Note

$$\frac{d\phi(a)}{da} = \theta(-a + c_S - c_A)h(c_I + \pi_I - a), \quad (8)$$

so  $\phi(a)$  obtains a local maximum<sup>11</sup> at  $a = a_T \equiv c_S - c_A$  (as illustrated in the figures). The value  $a_T$  **is conceptually the same** as what Rochet and Tirole (2008) call the “tourist test

---

<sup>11</sup>Indeed, formula (8) shows that  $\frac{d\phi}{da}$  is positive for  $a < a_T$  and negative for  $a > a_T$ , implying a local maximum of  $\phi$  for  $a = a_T$ .



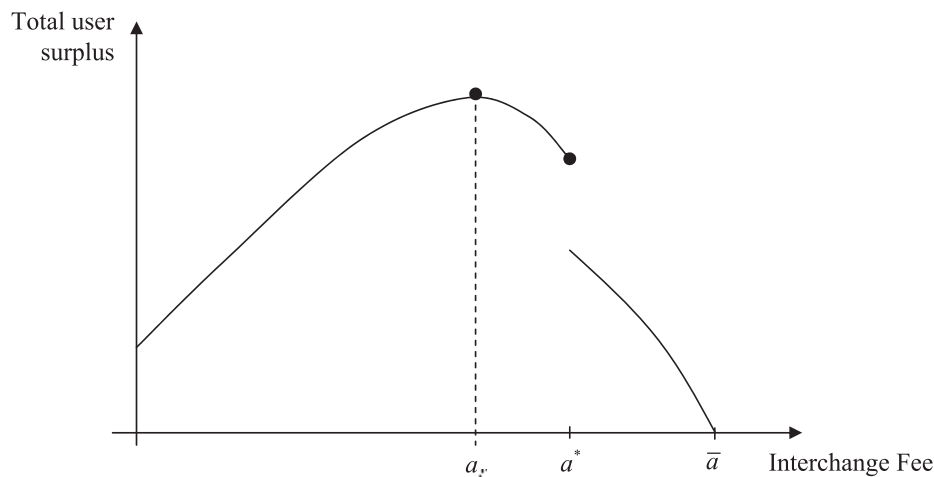
threshold”, i.e. the maximum level of the interchange fee such that the merchant fee  $m = c_A + a$  is less than the cost  $c_S$  of the relevant alternative technology (i.e. which in a credit card setting is store credit) for the merchant. This **threshold is also conceptually equivalent** to the interchange fee recommended by Baxter (1983) in the context of a perfectly competitive banking sector, and to Farrell’s (2006) “Merchant Indifference Criterion”. **Note, however**,  $a_T$  is not necessarily a global maximum for total user surplus  $\phi$ . This is because  $\phi$  has a downward jump at  $a = a^* \equiv c_I + \pi_I$ . This jump is due to the fact that when  $a > a^*$ , cardholder fee  $f$  becomes negative, and cardholders find it convenient to use their credit card for ordinary purchases. **In formula (6)** this is captured by the fact that  $L_c = 1$  **if and only if**  $a > a^*$ , implying a downward jump of magnitude  $c + \pi_I$  for total user surplus  $\phi$ . This jump complicates the analysis of consumer surplus maximisation, and introduces three regimes.

**In regime 1**,  $a_T = c_S - c_A$  is less than  $a^* \equiv c_I + \pi_I$ , and therefore total user surplus is maximized for  $a = a_T$ . This corresponds to the situation where credit card transactions are more costly to provide than retailer provided store credit:

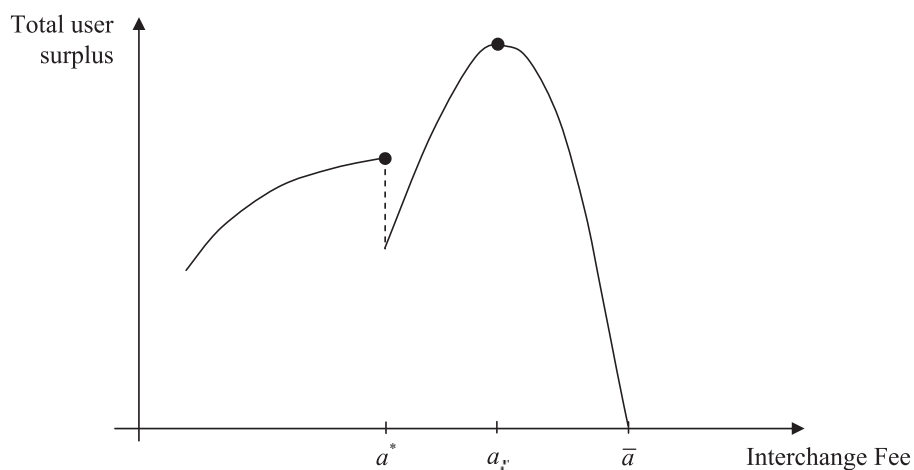
$$c_S - c_A \leq c_I + \pi_I \quad \Leftrightarrow \quad c_S \leq c + \pi_I.$$

In regime 1, regulation to maximize consumer surplus would require making sure credit cards were priced to be *more* expensive for consumers than using cash (or **debit cards**).

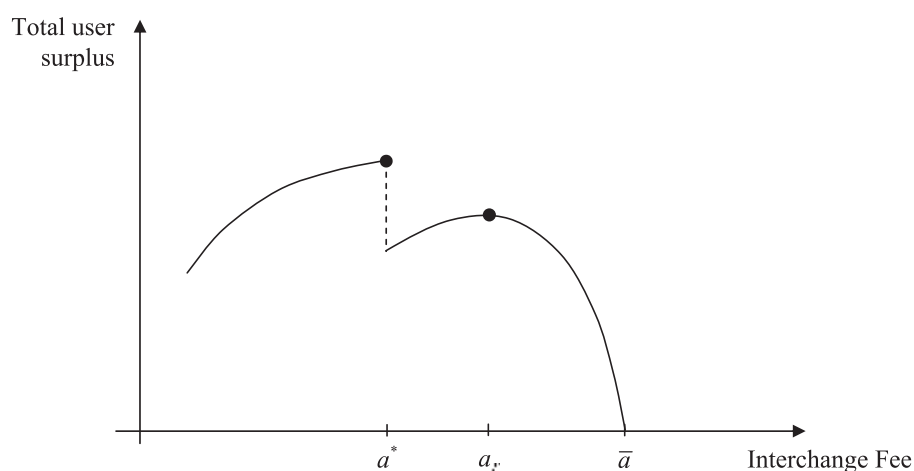
**The more realistic** cases are captured by regimes 2 and 3, which correspond to the reverse situation. In regime 2 the incidence of convenience users is not so large (so that  $a_T$  is still the maximum of total user surplus). By contrast in regime 3, the maximum of  $\phi$  is obtained at  $a^*$ .



First regime: total user surplus is maximum for  $a = a_T < a^*$ .



Second regime: total user surplus is maximum for  $a = a_T > a^*$ .



Third regime: total user surplus is maximum for  $a = a^*$ .

In all regimes,  $\bar{a}$  denotes the interchange fee selected by a monopoly network.

### 3.3 Policy Implications

Recall that the volume of credit card transactions, and thus the profit of banks, increase with  $a$ . This implies that a monopoly card network will set an interchange fee equal to the maximum level  $\bar{a}$  that is compatible with merchant acceptance of payment cards. This level is always greater than both  $a_T$  and  $a^*$ . Therefore, if competition authorities aim at maximizing (short-term) consumer surplus (or equivalently  $\phi(a)$ ), they will always find the privately optimal interchange

fee excessive.<sup>12</sup> A regulatory cap on interchange fees can therefore be justified but the appropriate level of the cap is not always the same: it is  $a_T$  in regime 2 and  $a^*$  in regime 3. Which of these regimes is relevant depends on the magnitude of the difference between  $c_S$ , the retailer's cost of providing store credit and  $c + \pi_I$ , the total user cost of credit transactions. We denote this difference by  $\delta \equiv c_S - c - \pi_I$ . These results are recapitulated in the next proposition:

**Proposition 3** *If regulatory authorities aim at maximizing (short-term) consumer surplus, privately optimal interchange fees are too high. A regulatory cap on interchange fees can therefore raise (short-term) consumer surplus, but two cases must be considered, depending on the value of  $\delta = c_S - c - \pi_I$ :*

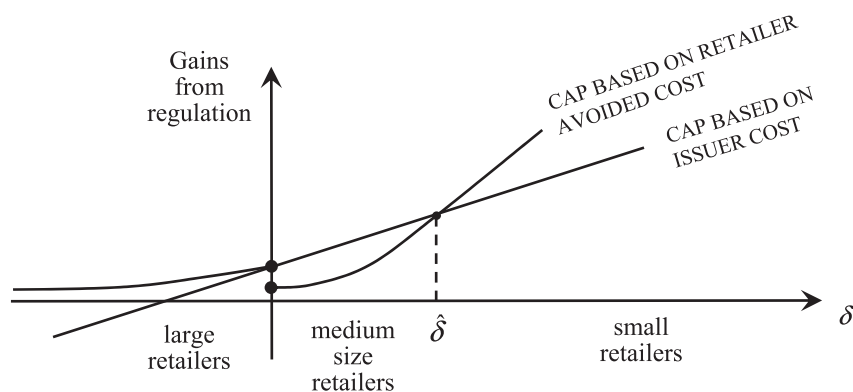
*a) If  $\delta < 0$  (regime 1) or if  $c + \pi_I \leq \theta \int_{-\delta}^0 (c_B + \delta) dH(c_B)$  (regime 2), the regulatory cap should be  $a_T = c_S - c_A$ .*

*b) Otherwise (regime 3), the regulatory cap should be  $a^* = c_I + \pi_I$ .*

**Proof of Proposition 3:** See the Appendix.

The proposition implies that the interchange fee  $a_T = c_S - c_A$  achieves the global maximum of  $\phi(a)$  provided (i)  $\delta < 0$  or equivalently  $a_T \leq a^* = c_I + \pi_I$  (i.e provided  $L_c = 0$  at  $a_T$ , which corresponds to regime 1) or (ii) if  $\delta > 0$  (i.e.  $L_c = 1$  at  $a_T$ ) provided that  $\theta \int_{-\delta}^0 (c_B + \delta) dH(c_B) \geq c + \pi_I$  (this corresponds to regime 2). Regime 2 arises when the drop in surplus caused by inefficient use of credit cards for ordinary purchases is not too great. **Regime 3 arises** when the deadweight loss from excessive use of credit cards by convenience users dominates. In either case, lowering interchange fees from the private maximum to  $a_T$  unambiguously raises consumer surplus. This is immediately deduced from the following figure, which plots the gains from regulation as a function of the incremental cost parameter  $\delta$ , which can be expected to be smaller for large retailers than for small retailers. The figure shows that a cap based on retailer avoided cost ( $a \leq a_T$ ) always increases consumer surplus, while a cap based on issuer cost ( $a \leq a^*$ ) may sometimes decrease consumer surplus (when  $\delta$  is negative enough) with respect to the unregulated monopoly situation, **as may be the case for very large retailers.**

<sup>12</sup>This is not necessarily true anymore if competition authorities aim at maximizing long-term consumer surplus or social welfare, which includes banks' profits (see Rochet and Tirole, 2002, 2008) or if merchants are heterogeneous (see Wright, 2004).



Therefore, given it would be extremely difficult in practice for a regulator to determine which **of these regimes prevails**, a conservative approach to regulation would be to cap the interchange fee at  $a_T$ . This cap is relatively easy to construct. It is based on a retailer's net avoided costs. This is the retailer's cost saving from not having to provide credit directly, or in other words, the cost of providing credit itself less the cost of the acquiring service offered by its bank (i.e. the costs incurred by the retailer's bank in providing the retailer with the ability to accept credit cards, including any such costs that are passed onto the merchant other than the interchange fee).<sup>13</sup> In contrast, the interchange fee  $a^*$  (which is based on the issuers' cost) represents a lower bound for desirable interchange fees **if either of the more realistic regimes 2 or 3 hold, since then** any interchange fee lower than  $a^*$  unambiguously lowers consumers surplus and welfare. This could be used as a check to make sure the calculated  $a_T$  is not too low. Thus, regulations which base interchange fees on **issuer costs could be rationalized** in our framework, although **realistically, only as a lower bound** on what might be desirable for consumers.

## 4 Alternative specifications

In reaching the policy implications above, we adopted a model in which cardholders sometimes prefer using store credit for credit purchases. Despite this, in the model consumers never use store cards for ordinary purchases. This reflected our assumption that the costs to consumers and retailers of using credit cards and store credit (and therefore the possibility of a net benefit of store credit to consumers) **only arises with respect to the use of credit**. According to this view, consumers

<sup>13</sup>Note that  $a_T$  may differ for different classes of retailers by contrast with  $a^*$  which only depends on the issuer.

never have an incentive to use store credit for ordinary purchases given credit is not needed for these purchases. An alternative view is that the costs of store credit measure the costs of using these instruments for transactions in general, regardless of whether they are ordinary purchases or credit purchases. Consistent with this view, in this section, we consider two different settings in which we assume the costs of using each type of payment are the same across ordinary and credit purchases. In the first setting the cost to consumers of using store credit is allowed to take negative values. The idea is that store credit may sometimes provide benefits to consumers even for ordinary purchases, such as possibly allowing cardholders to preserve their funds for some other contingencies. Thus, consumers will choose between cash, store credit, and credit cards (if they hold credit cards) when making ordinary purchases. This change in assumptions turns out only to strengthen our main results. Second, we consider the case in which the cost of using store credit is always positive ( $c_B \geq 0$ ) so that consumers would never want to use store credit for ordinary purchases (cash would always be preferred). Instead, we endogenize the choice by consumers of whether to hold a credit card in the first place. We show this framework still leads to similar policy implications to our benchmark case.

#### 4.1 Allowing store credit for ordinary purchases

We follow the same steps as in section 3. The main impact of the additional types of possible payment transactions is to make the definitions of the underlying terms  $M_i$ ,  $\Gamma(a)$ ,  $S(a)$  and  $\phi(a)$  more complicated, while leaving the structure of the analysis largely unaffected. As we will see, allowing consumers to choose store credit when making ordinary purchases only strengthens the policy implications of section 3. Essentially, it increases the number of situations in which it is desirable to have consumers internalize the cost to merchants of providing store credit, and thereby makes it more likely that the interchange fee  $a_T = c_S - c_A$  maximizes consumer surplus.

We start by considering the retailer's margins and demand arising from each type of transaction. A fraction  $1 - xL_i$  of the time, consumers cannot use credit cards. For ordinary purchases, these consumers will use store credit if  $c_B \leq 0$ , which happens with probability  $H(0)$ ; otherwise, with probability  $1 - H(0)$  they will prefer to use cash. For credit purchases, which happen with probability  $\theta$ , these consumers will always use store credit. A fraction  $xL_i$  of the time, consumers can also use credit cards. For ordinary purchases, these cardholders will use store credit if  $c_B < L_c f$ , and otherwise either credit cards (if  $L_c = 1$ ) or cash (if  $L_c = 0$ ). For credit

purchases, these cardholders will use store credit if  $c_B < f$  and otherwise credit cards. Collecting together a retailer's margins associated with each of these different possibilities, retailer  $i$ 's expected margin per-customer is therefore equal to

$$\begin{aligned} M_i = & (1 - xL_i) [(H(0) + \theta)(p_i - \gamma - c_S) + (1 - H(0))(p_i - \gamma)] \\ & + xL_i \left[ H(L_c f)(p_i - \gamma - c_S) + (1 - H(L_c f))(p_i - \gamma - L_c m) \right. \\ & \left. + \theta(H(f))(p_i - \gamma - c_S) + (1 - H(f))(p_i - \gamma - m) \right] \end{aligned}$$

or after simplifying

$$M_i = (1 + \theta)(p_i - \gamma) - (H(0) + \theta)c_S - x\Gamma(a)L_i,$$

where

$$\Gamma(a) = (H(L_c f) - H(0))c_S + (1 - H(L_c f))L_c m + \theta(1 - H(f))(m - c_S).$$

Similarly, we can rewrite consumers' expected utility and retailer  $i$ 's market share taking into account the additional types of payment transactions that are possible. Retailer  $i$  offers an expected surplus (ignoring transportation costs) of

$$\begin{aligned} U_i = & (1 - xL_i) \left( u_0 + \theta u_1 - (1 + \theta)p_i - \int_{\underline{c}_B}^0 c_B dH(c_B) - \theta E(c_B) \right) \\ & + xL_i \left( u_0 + \theta u_1 - (1 + \theta)p_i - \int_{\underline{c}_B}^{L_c f} c_B dH(c_B) - \int_{L_c f}^{\bar{c}_B} L_c f dH(c_B) - \theta \left( \int_{\underline{c}_B}^f c_B dH(c_B) + \int_f^{\bar{c}_B} f \right) \right) \end{aligned}$$

or after simplifying

$$U_i = u_0 + \theta u_1 - (1 + \theta)p_i - \int_{\underline{c}_B}^0 c_B dH(c_B) - \theta E(c_B) + xS(a)L_i,$$

where

$$S(a) = \int_{L_c f}^0 (c_B - L_c f) dH(c_B) - \int_0^{\bar{c}_B} L_c f dH(c_B) + \theta \int_f^{\bar{c}_B} (c_B - f) dH(c_B).$$

Total user surplus can then be written

$$\begin{aligned} \phi(a) = & S(a) - \Gamma(a) \\ = & \int_{L_c(c_I + \pi_I - a)}^0 (c_B + c_S - (c + \pi_I)L_c) dH(c_B) \\ & + \theta \int_{c_I + \pi_I - a}^{\bar{c}_B} (c_B + c_S - c - \pi_I) dH(c_B) - \int_0^{\bar{c}_B} (c + \pi_I)L_c dH(c_B). \end{aligned}$$

The last two terms in  $\phi(a)$  are similar to before. The first term is new, reflecting that there may be cost savings from using credit cards for ordinary purchases if this avoids the use of more costly



store credit. Together, the first two terms represent total expected cost savings of the two users (consumers and retailers) from using credit cards for consumers' purchases (as opposed to using store credit). The last term is the deadweight loss associated with convenience usage of credit cards for ordinary purchases, although this now arises for fewer transactions since store credit is now used for some ordinary transactions.

Aside from these new surplus definitions, the analysis of retailers' decisions remains unchanged. That is, retailer  $i$ 's market share and profit expressions remain the same, as do propositions 1 and 2. Retailer  $i$ 's equilibrium price is now given by

$$(1 + \theta) p_i^* = t + \gamma(1 + \theta) + (H(0) + \theta) c_S + x\Gamma(a)L_i + \frac{x}{3}\phi(a)(L_i - L_j),$$

where the right-hand-side includes the additional term  $H(0)c_S$  arising from the use of store credit for some ordinary purchases by cash customers.<sup>14</sup> Consumer surplus is similarly adjusted for this additional term, and so now equals

$$CS(a) = u_0 + \theta u_1 - \gamma(1 + \theta) - \frac{5t}{4} - \int_{c_B}^0 (c_B + c_S) dH(c_B) - \theta(E(c_B) + c_S) + x\phi(a).$$

Thus, aggregate consumer surplus is still equal, up to an additive and a positive multiplicative constants, to total user surplus  $\phi(a)$ .

As before  $\phi(a)$  has a jump (down) at  $a = a^*$ . However, the total user surplus function is now more likely to be maximized for a higher interchange fee. The derivative  $\phi'(a)$  is the same as before except it is multiplied by  $(L_c + \theta)/\theta$  which exceeds one for  $a > a^*$ . This also implies  $a_T$  which maximizes  $\phi(a)$  for  $a > a^*$  is identical to before. That is, as before  $a_T = c_S - c_A$ . Proposition 3 now becomes

**Proposition 4** *If regulatory authorities aim at maximizing (short-term) consumer surplus, privately optimal interchange fees are too high. A regulatory cap on interchange fees can therefore raise (short-term) consumer surplus, but two cases must be considered:*

- a) If  $\delta < 0$  (**regime 1**) or if  $(c + \pi_I)(1 - H(0)) \leq (1 + \theta) \int_{-\delta}^0 (c_B + \delta) dH(c_B)$  (**regime 2**), the regulatory cap should be  $a_T = c_S - c_A$ .*
- b) Otherwise (**regime 3**), the regulatory cap should be  $a^* = c_I + \pi_I$ .*

<sup>14</sup>It is no longer the case that  $\Gamma(a)$  is necessarily increasing in  $a$  since  $\Gamma'(a) = (L_c + \theta)(h(f)(m - c_S) + (1 - H(f)))$  so that for  $m$  much below  $c_S$ , higher interchange fees could actually lower prices. However, for  $m$  not too much below  $c_S$  (i.e.  $a$  not too much below  $c_S - c_A$ ), then prices will still be increasing in interchange fees.

The proposition implies that the interchange fee  $a_T = c_S - c_A$  achieves the global maximum of  $\phi(a)$  provided (i)  $\delta < 0$ , or equivalently  $a_T \leq a^* = c_I + \pi_I$  (i.e. provided  $L_c = 0$  at  $a_T$ ) or (ii) if  $\delta > 0$  (i.e.  $L_c = 1$  at  $a_T$ ) provided

$$(1 + \theta) \int_{-\delta}^0 (c_B + \delta) dH(c_B) \geq (c + \pi_I) (1 - H(0)). \quad (9)$$

Notice the trade-off here. The left-hand-side of (9) measures the cost saving from using credit cards rather than store credit when credit cards have a negative fee as opposed to a zero fee. The right-hand-side of (9) measures the additional costs to users from the use of credit cards rather than cash for ordinary purchases when credit cards have a negative fee (and so are used whenever  $c_B > 0$ ). Compared to before, the condition for case (ii) to apply is more easily satisfied since the left-hand-side integral is the same as before except it is multiplied by  $1 + \theta$  instead of  $\theta$  and the right-hand-side integral is the same as before except it is multiplied by  $1 - H(0)$  instead of 1. In other words, allowing for consumers to choose store credit when making ordinary purchases only strengthens the previous policy implications.

## 4.2 Endogenizing card membership decisions

In this section, we consider an extension of the existing model in which consumers make prior membership decisions, on whether to hold a credit card or not. Prior membership decisions are potentially important since if interchange fee are set too low (for instance, at zero), then consumers can be expected to face higher fees (for instance, the full cost of issuing cards) at which point many consumers may choose to no longer hold credit cards. In order to simplify the resulting model, we assume that consumers always view using store credit as costly ( $\underline{c}_B = 0$ ). With this assumption there is no issue about whether consumers can use store credit for ordinary purchases; even if they could use store credit for these purchases, as was the case in section 4.1, they would always prefer to use cash for this purpose. This assumption on  $c_B$  will also imply that, whenever  $f < 0$ , cardholders will *always* want to use cards if they are accepted. In this case, cardholders will never use store credit. However, as we will see, setting an interchange fee such that  $f < 0$  may still be optimal for consumers (as opposed to setting an interchange fee such that  $f = 0$ ) since it induces more people to hold a credit card, thereby reducing the use of expensive store credit for credit purchases by consumers who otherwise would not hold a credit card.

The model is the same as before except that (i) there is an additional choice for consumers as to whether to hold a card or not and (ii) we set  $\underline{c}_B = 0$ . Specifically, at the same time as

retailers choose their card acceptance policies (i.e. stage 3), consumers receive a random draw  $v$  of the benefit of holding a credit card and decide whether to hold the card or not.  $v$  is drawn according to the distribution function  $\Psi$  over the support  $(-\infty, \bar{v}]$  for some  $\bar{v} > 0$ . Let  $x(a)$  be the measure of consumers who join given an interchange fee  $a$ . Some consumers draw  $v < 0$ , so they would prefer not to hold a credit card other things equal (e.g. if they didn't expect to use it). This could also capture that there are significant per-customer costs to issuers associated with managing a cardholder which are passed through to cardholders. For other consumers, cards may offer more than just the ability to make transactions at retailers, in which case  $v > 0$  is possible.

We follow the same steps as in section 3. The assumption  $c_B \geq 0$  is consistent with the analysis in section 3 which assumes consumers never use store credit for ordinary purchases. Therefore, the analysis of retailers' decisions (which treat  $x$  as given), and the various surplus expressions are identical to before. Note, however, that when  $f < 0$ , then  $1 - H(f) = 1$  since all cardholders will use credit cards rather than store credit, which is just a special case of the more general analysis detailed in section 3. This implies

$$\phi(a) = \theta \int_{(1-L_c)f}^{\bar{c}_B} (c_B + c_S - c - \pi_I) dH(c_B) - (c + \pi_I)L_c,$$

which is now independent of  $a$  if  $a > a^*$  (so that  $L_c = 1$ ):

$$a > a^* \Rightarrow \phi(a) \equiv \theta[E(c_B) + c_S] - (1 + \theta)(c + \pi_I).$$

In order to keep the analysis interesting, we assume

$$\theta(E(c_B) + c_S) > (1 + \theta)(c + \pi_I), \quad (10)$$

so that the surplus created from credit cards (the cost saving) is higher than their additional cost to users when they are always used (i.e. when  $f < 0$ ). Without making the assumption in (10), retailers would reject cards if and only if  $a > a^*$ . Taking this into account, the card network would set the interchange fee at  $a^*$  and there would be no rationale for any regulation. With (10), the card network will want to set  $a > a^*$ , after which the interchange fee is neutral in terms of usage decisions. Retailers will always accept cards no matter how high the fees and cardholders will always use them. From the point of view of users (in aggregate), interchange fees beyond  $a^*$  just represent pure transfer fees. It will therefore be optimal for the card scheme to set interchange fees to the point where all consumers hold cards. Not surprisingly, like before, this will lead interchange fees to be set too high.

Using the expressions for  $U_{\text{credit}}(a)$  and  $U_{\text{cash}}(a)$  from section 3 but taking into account that  $c_B = 0$ , the expected additional utility to a consumer of holding a card is

$$U_{\text{credit}}(a) - U_{\text{cash}}(a) = \theta \int_{(1-L_c)f}^{\bar{c}_B} (c_B - f) dH(c_B) - L_c f$$

so consumers will hold cards if

$$v > L_c f - \theta \int_{(1-L_c)f}^{\bar{c}_B} (c_B - f) dH(c_B)$$

implying  $x(a) = 1 - \Psi \left( L_c f - \theta \int_{(1-L_c)f}^{\bar{c}_B} (c_B - f) dH(c_B) \right)$  which is increasing in  $a$ . Taking into account the additional utility from holding a card, total consumer surplus is equal to (up to an additive and a multiplicative constants)

$$\int_{L_c f - \theta \int_{(1-L_c)f}^{\bar{c}_B} (c_B - f) dH(c_B)}^{\bar{v}} (v + \phi(a)) d\Psi(v).$$

When  $a > a^*$ , this simplifies to

$$\int_{(1+\theta)f - \theta E(c_B)}^{\bar{v}} (v + \theta(E(c_B) + c_S) - (1 + \theta)(c + \pi_I)) d\Psi(v).$$

As before, total surplus has a jump (down) at  $a = a^*$ . However, the total surplus function behaves slightly differently for  $a > a^*$ . It is maximum when

$$(1 + \theta)f - \theta E(c_B) = (1 + \theta)(c + \pi_I) - \theta[E(c_B) + c_S].$$

After simplifying, we obtain the optimal value of  $f$ :

$$f = c + \pi_I - \frac{\theta}{1 + \theta} c_S.$$

Since  $f$  is equal to  $c_I + \pi_I - a$ , we obtain that the interchange fee  $a_T$  which maximizes total user surplus for  $a > a^*$  is now

$$a_T = \frac{\theta}{1 + \theta} c_S - c_A. \quad (11)$$

This is lower than  $a_T = c_S - c_A$  in the benchmark model. Here the purpose of the optimal interchange fee is to induce consumers to hold cards even when they otherwise would not want to, so that they internalize retailers' surplus from being able to accept their credit cards rather than rely on store credit for credit purchases. Note the retailers' surplus  $c_S - c_A$  only arises a fraction  $\theta/(1 + \theta)$  of the time, whereas a fraction  $1 - \theta/(1 + \theta)$  of the time, the retailer is actually worse off by  $c_A$  due to the excessive usage of cards. Following this interpretation,  $a_T$  can also be written:

$$a_T = \frac{\theta}{1 + \theta} (c_S - c_A) - \left(1 - \frac{\theta}{1 + \theta}\right) c_A.$$

As in Proposition 3, which of  $a_T$  or  $a^*$  maximizes consumer surplus depends on the magnitude of the incremental cost of store cards over credit cards, which is now measured by

$$\hat{\delta} = \theta c_S - (1 + \theta)(c + \pi_I).$$

**Proposition 5** *If regulatory authorities aim at maximizing (short-term) consumer surplus, privately optimal interchange fees are too high. A regulatory cap on interchange fees can therefore raise (short-term) consumer surplus, but two cases must be considered:*

*a) If  $\hat{\delta} < 0$  (regime 1) or if*

$$\int_{-\theta E(c_B)}^{\bar{v}} (c + \pi_I) d\Psi(v) \leq \int_{-\hat{\delta} - \theta E(c_B)}^{-\theta E(c_B)} \left( v + \hat{\delta} + \theta E(c_B) \right) d\Psi(v) \quad (12)$$

*(regime 2), the regulatory cap should be  $a_T = \frac{\theta}{1+\theta} c_S - c_A$ .*

*b) Otherwise (regime 3), the regulatory cap should be  $a^* = c_I + \pi_I$ .*

The proposition implies that the interchange fee  $a_T = \frac{\theta}{1+\theta} c_S - c_A$  achieves the global maximum of  $\phi(a)$  provided (i)  $\hat{\delta} < 0$ , or equivalently  $a_T \leq a^* = c_I + \pi_I$  (i.e. provided  $L_c = 0$  at  $a_T$ ) or (ii) if  $a_T > a^*$  (i.e.  $L_c = 1$  at  $a_T$ ) provided (12) holds. Given it may be hard to measure  $\theta$  with any confidence or to evaluate whether (12) holds or not, a conservative regulatory approach could again be to use  $a_T = c_S - c_A$  as the regulatory cap since according to the model this (i) increases consumer surplus relative to the privately set interchange fee; (ii) possibly maximizes consumer surplus; and (iii) yet is never too low from the point of view of maximizing (short-term) consumer surplus. Thus, in this alternative framework, policy implications remain remarkably similar to our benchmark case in section 3.

## 5 Conclusions

Much of the existing literature on interchange fees treats payment cards as though they were debit cards. This paper provides a new theory of interchange fees that is specifically applicable to credit card networks. The model we provide captures a trade-off that can arise between the excessive usage of credit cards for ordinary purchases and the importance of getting cardholders who do need credit to internalize retailers' avoided costs arising from their credit card usage. In terms of this trade-off, we find that an unregulated card network always sets the interchange fee too high. Consumer surplus can be increased by imposing a cap on interchange fees which equals

the retailers' net avoided costs from not having to provide credit themselves. Further lowering interchange fees from this level towards issuing costs may either increase or decrease consumer surplus, although initially consumer surplus always falls in our framework.

The possibility that the regulator's preferred interchange fee predicted by our theory may still be moderately high reflects the potentially high merchant benefits of accepting cards in these circumstances, which cardholders may not otherwise internalize. Put differently, if the interchange fee is set too low (say at zero) so that consumers were sometimes not willing to hold or use cards for such transactions, then competing merchants will instead tend to rely more on store credit (or other forms of credit) so as to attract business. Quite plausibly, the additional costs to society of making greater use of these more expensive forms of credit will outweigh any benefit from encouraging debit rather than credit card transactions for ordinary purchases. Some excessive use of credit cards may be unavoidable given merchants cannot easily observe if credit is needed or not by their customers. This seems no different from the fact merchants that offer interest-free installment plans to their customers, will sometimes (perhaps often) end up offering these plans to consumers who actually do not need them.

One important direction for future research is to extend our model to allow retailers to offer different prices (through the use of discounts, interest-free periods and rewards) when consumers make use of store credit. In not allowing this possibility, we had in mind such price differentials being exogenous for each retailer, determined perhaps through an association of retailers or being absent altogether due to the inability of retailers to set differential prices based on a consumer's choice of payment technology. If retailers could discriminate based on the use of store credit they may be able to induce consumers to use credit cards and store credit efficiently. However, any individual retailer would still not be able to (nor have the incentive to) use its differential pricing to get consumers to make the right decision about whether to hold a credit card in the first place. As such, our results from section 4, in which card membership is endogenized, are likely to carry over to the case retailers can price discriminate in this way.

Another possible direction for future research is to extend our model to allow for competing payment networks. By adapting the arguments of Guthrie and Wright (2007) one should be able to show similar results to those shown here still hold



when there is competition between multiple card payment networks. The idea is that competing networks will seek to choose an interchange fee somewhere between the one maximizing total user surplus and that chosen by a monopoly network, depending on the extent to which consumers choose to hold multiple cards, and so these may still be too high.

We conclude by noting some other implications of our theory, which may be able to explain real-world observations that have previously defied theoretical explanation. If credit is more likely to be needed by customers for large purchases, then the optimal interchange fee should be ad valorem in our setting (thereby better targeting the transfer to cardholders for the types of transactions where credit is needed). Thus, the model potentially provides a justification for the widespread use of ad valorem credit card interchange fees. It also explains why merchants may want to reject credit cards for small transactions (where people are more likely to be able to purchase anyway using cash). Most importantly, it explains why interchange fees are typically lower for debit cards than for credit cards. Finally, the theory suggests large retailers that are able to gain a competitive advantage over smaller rivals from being able to offer their own store-credit to customers, may have an interest in opposing the widespread use of general purpose credit cards.

## Appendix

### Proof of Proposition 1:

From the text we have

$$\pi_i = ((1 + \theta)(p_i - \gamma) - \theta c_S - x\Gamma(a)L_i) \left( \frac{1}{2} + \frac{(1 + \theta)(p_j - p_i)}{2t} + xS(a) \frac{L_i - L_j}{2t} \right)$$

Differentiating we get

$$\frac{2t}{1 + \theta} \frac{\partial \pi_i}{\partial p_i} = t + (1 + \theta)(p_j - p_i) + xS(a)(L_i - L_j) - (1 + \theta)(p_i - \gamma) + \theta c_S + xL_i\Gamma(a).$$

At a Nash equilibrium, we have for  $i, j = 1, 2$ :

$$(1 + \theta)(2p_i - p_j) = t + \gamma(1 + \theta) + \theta c_S + x(S(a)(L_i - L_j) + L_i\Gamma(a)).$$

Solving for  $p_i$ , we obtain formula (5):

$$(1 + \theta)p_i = t + \gamma(1 + \theta) + \theta c_S + x\Gamma(a)L_i + \frac{x}{3}\phi(a)(L_i - L_j). \quad (13)$$

Substituting (13) into (3) and using (4) implies formula (6).

### Proof of Proposition 3:

It results immediately from the comparison of  $\phi(a_T)$  and  $\phi(a^*)$ . Recall the expression of  $\phi(a)$ :

$$\phi(a) = \theta \int_{a^* - a}^{\infty} (c_B + c_S - c - \pi_I) dH(c_B) - (c + \pi_I) \mathbb{I}_{a > a^*},$$

and note that  $\delta = c_S - c_A - c_I - \pi_I = a_T - a^*$ . Thus

$$\phi(a_T) = \theta \int_{-\delta}^{\infty} (c_B + \delta) dH(c_B) - (c + \pi_I) \mathbb{I}_{\delta > 0},$$

while

$$\phi(a^*) = \theta \int_0^{\infty} (c_B + \delta) dH(c_B).$$

Thus

$$\phi(a_T) - \phi(a^*) = \theta \int_{-\delta}^0 (c_B + \delta) dH(c_B) - (c + \pi_I) \mathbb{I}_{\delta > 0}.$$

When  $\delta < 0$  we have

$$\phi(a_T) - \phi(a^*) = -\theta \int_0^{-\delta} (c_B + \delta) dH(c_B) > 0$$

(since  $c_B + \delta \leq 0$  when  $c_B$  belongs to  $[0, -\delta]$ ). Thus  $\phi$  is maximum for  $a = a_T$ .

When  $\delta > 0$  we have

$$\phi(a_T) - \phi(a^*) = \theta \int_{-\delta}^0 (c_B + \delta) dH(c_B) - (c + \pi_I).$$

This establishes Proposition 3.

## References

- BAXTER, W.P. (1983) "Bank Interchange of Transactional Paper: Legal Perspectives," *Journal of Law and Economics*, 26: 541-88.
- BOLT, W. AND S. CHAKRAVORTY (2008) "Consumer Choice and Merchant Acceptance of Payment Media," mimeo.
- BRADFORD, T. AND F. HAYASHI (2008) "Developments in Interchange Fees in the United States and Abroad," Payments System Research Briefing, Federal Reserve Bank of Kansas City. April.
- CHAKRAVORTY, S. AND T. TO (2007) "A Theory of Credit Cards," *International Journal of Industrial Organization*, 25: 583-95.
- FARRELL, J. (2006) "Efficiency and Competition between Payment Instruments," *Review of Network Economics*, 5: 26-44.
- GUTHRIE, G. AND J. WRIGHT (2007) "Competing Payment Schemes," *Journal of Industrial Economics*, 55: 37-67.
- ROCHET, J.-C. (2003) "The Theory of Interchange Fees: A Synthesis of Recent Contributions," *Review of Network Economics*, 2: 97-124.
- ROCHET, J.-C. AND J. TIROLE (2002) "Cooperation among Competitors: Some Economics of Payment Card Associations," *Rand Journal of Economics*, 33: 549-70.
- ROCHET, J.-C. AND J. TIROLE (2008) "Must-Take Cards: Merchant Discounts and Avoided Costs," mimeo, Toulouse School of Economics.
- SCHMALENSEE, R. (2002) "Payment Systems and Interchange Fees," *Journal of Industrial Economics*, 50: 103-22.
- WEINER, S. AND J. WRIGHT (2005) "Interchange Fees in Various Countries: Developments and Determinants," *Review of Network Economics*, 4: 290-323.
- WRIGHT, J. (2003) "Optimal Card Payment Systems," *European Economic Review*, 47: 587-612.
- WRIGHT, J. (2004) "The Determinants of Optimal Interchange Fees in Payment Systems," *Journal of Industrial Economics*, 52: 1-26.

## European Central Bank Working Paper Series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (<http://www.ecb.europa.eu>).

- 1086 “Euro area money demand: empirical evidence on the role of equity and labour markets” by G. J. de Bondt, September 2009.
- 1087 “Modelling global trade flows: results from a GVAR model” by M. Bussière, A. Chudik and G. Sestieri, September 2009.
- 1088 “Inflation perceptions and expectations in the euro area: the role of news” by C. Badarınza and M. Buchmann, September 2009.
- 1089 “The effects of monetary policy on unemployment dynamics under model uncertainty: evidence from the US and the euro area” by C. Altavilla and M. Ciccarelli, September 2009.
- 1090 “New Keynesian versus old Keynesian government spending multipliers” by J. F. Cogan, T. Cwik, J. B. Taylor and V. Wieland, September 2009.
- 1091 “Money talks” by M. Hoerova, C. Monnet and T. Temzelides, September 2009.
- 1092 “Inflation and output volatility under asymmetric incomplete information” by G. Carboni and M. Ellison, September 2009.
- 1093 “Determinants of government bond spreads in new EU countries” by I. Alexopoulou, I. Bunda and A. Ferrando, September 2009.
- 1094 “Signals from housing and lending booms” by I. Bunda and M. Ca’Zorzi, September 2009.
- 1095 “Memories of high inflation” by M. Ehrmann and P. Tzamourani, September 2009.
- 1096 “The determinants of bank capital structure” by R. Gropp and F. Heider, September 2009.
- 1097 “Monetary and fiscal policy aspects of indirect tax changes in a monetary union” by A. Lipińska and L. von Thadden, October 2009.
- 1098 “Gauging the effectiveness of quantitative forward guidance: evidence from three inflation targeters” by M. Andersson and B. Hofmann, October 2009.
- 1099 “Public and private sector wages interactions in a general equilibrium model” by G. Fernández de Córdoba, J.J. Pérez and J. L. Torres, October 2009.
- 1100 “Weak and strong cross section dependence and estimation of large panels” by A. Chudik, M. Hashem Pesaran and E. Tosetti, October 2009.
- 1101 “Fiscal variables and bond spreads – evidence from eastern European countries and Turkey” by C. Nickel, P. C. Rother and J. C. Rülke, October 2009.
- 1102 “Wage-setting behaviour in France: additional evidence from an ad-hoc survey” by J. Montornés and J.-B. Sauner-Leroy, October 2009.
- 1103 “Inter-industry wage differentials: how much does rent sharing matter?” by P. Du Caju, F. Rycx and I. Tojerow, October 2009.

- I 104 “Pass-through of external shocks along the pricing chain: a panel estimation approach for the euro area” by B. Landau and F. Skudelny, November 2009.
- I 105 “Downward nominal and real wage rigidity: survey evidence from European firms” by J. Babecký, P. Du Caju, T. Kosma, M. Lawless, J. Messina and T. Rõõm, November 2009.
- I 106 “The margins of labour cost adjustment: survey evidence from European firms” by J. Babecký, P. Du Caju, T. Kosma, M. Lawless, J. Messina and T. Rõõm, November 2009.
- I 107 “Interbank lending, credit risk premia and collateral” by F. Heider and M. Hoerova, November 2009.
- I 108 “The role of financial variables in predicting economic activity” by R. Espinoza, F. Fornari and M. J. Lombardi, November 2009.
- I 109 “What triggers prolonged inflation regimes? A historical analysis.” by I. Vansteenkiste, November 2009.
- I 110 “Putting the New Keynesian DSGE model to the real-time forecasting test” by M. Kolasa, M. Rubaszek and P. Skrzypczyński, November 2009.
- I 111 “A stable model for euro area money demand: revisiting the role of wealth” by A. Beyer, November 2009.
- I 112 “Risk spillover among hedge funds: the role of redemptions and fund failures” by B. Klaus and B. Rzepkowski, November 2009.
- I 113 “Volatility spillovers and contagion from mature to emerging stock markets” by J. Beirne, G. M. Caporale, M. Schulze-Ghattas and N. Spagnolo, November 2009.
- I 114 “Explaining government revenue windfalls and shortfalls: an analysis for selected EU countries” by R. Morris, C. Rodrigues Braz, F. de Castro, S. Jonk, J. Kremer, S. Linehan, M. Rosaria Marino, C. Schalck and O. Tkacevs.
- I 115 “Estimation and forecasting in large datasets with conditionally heteroskedastic dynamic common factors” by L. Alessi, M. Barigozzi and M. Capasso, November 2009.
- I 116 “Sectorial border effects in the European single market: an explanation through industrial concentration” by G. Cafiso, November 2009.
- I 117 “What drives personal consumption? The role of housing and financial wealth” by J. Slacalek, November 2009.
- I 118 “Discretionary fiscal policies over the cycle: new evidence based on the ESCB disaggregated approach” by L. Agnello and J. Cimadomo, November 2009.
- I 119 “Nonparametric hybrid Phillips curves based on subjective expectations: estimates for the euro area” by M. Buchmann, December 2009.
- I 120 “Exchange rate pass-through in central and eastern European member states” by J. Beirne and M. Bijsterbosch, December 2009.
- I 121 “Does finance bolster superstar companies? Banks, Venture Capital and firm size in local U.S. markets” by A. Popov, December 2009.
- I 122 “Monetary policy shocks and portfolio choice” by M. Fratzscher, C. Saborowski and R. Straub, December 2009.
- I 123 “Monetary policy and the financing of firms” by F. De Fiore, P. Teles and O. Tristani, December 2009.
- I 124 “Balance sheet interlinkages and macro-financial risk analysis in the euro area” by O. Castrén and I. K. Kavonius, December 2009.

- I 125 “Leading indicators in a globalised world” by F. Fichtner, R. Ruffer and B. Schnatz, December 2009.
- I 126 “Liquidity hoarding and interbank market spreads: the role of counterparty risk” by F. Heider, M. Hoerova and C. Holthausen, December 2009.
- I 127 “The Janus-headed salvation: sovereign and bank credit risk premia during 2008-09” by J. W. Ejsing and W. Lemke, December 2009.
- I 128 “EMU and the adjustment to asymmetric shocks: the case of Italy” by G. Amisano, N. Giammarioli and L. Stracca, December 2009.
- I 129 “Determinants of inflation and price level differentials across the euro area countries” by M. Andersson, K. Masuch and M. Schiffbauer, December 2009.
- I 130 “Monetary policy and potential output uncertainty: a quantitative assessment” by S. Delle Chiaie, December 2009.
- I 131 “What explains the surge in euro area sovereign spreads during the financial crisis of 2007-09?” by M.-G. Attinasi, C. Checherita and C. Nickel, December 2009.
- I 132 “A quarterly fiscal database for the euro area based on intra-annual fiscal information” by J. Paredes, D. J. Pedregal and J. J. Pérez, December 2009.
- I 133 “Fiscal policy shocks in the euro area and the US: an empirical assessment” by P. Burriel, F. de Castro, D. Garrote, E. Gordo, J. Paredes and J. J. Pérez, December 2009.
- I 134 “Would the Bundesbank have prevented the great inflation in the United States?” by L. Benati, December 2009.
- I 135 “Return to retail banking and payments” by I. Hasan, H. Schmiedel and L. Song, December 2009.
- I 136 “Payment scale economies, competition, and pricing” by D. B. Humphrey, December 2009.
- I 137 “Regulating two-sided markets: an empirical investigation” by S. Carbó-Valverde, S. Chakravorti and F. Rodríguez Fernández, December 2009.
- I 138 “Credit card interchange fees” by J.-C. Rochet and J. Wright, December 2009.



