Journal of Banking & Finance 34 (2010) 1788–1797

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Credit card interchange fees ☆

Jean-Charles Rochet [a], Julian Wright [b,*]

[a] Toulouse School of Economics, Université Toulouse 1 Capitole, Manufacture des Tabacs – Aile Jean-Jacques Laffont, 21 allée de Brienne, F-31000 Toulouse, France
[b] Department of Economics, National University of Singapore, 1 Arts Link, Singapore 117570, Singapore

ARTICLE INFO

Article history:
Received 16 February 2009
Accepted 28 February 2010
Available online 3 March 2010

JEL classification:
L11
E42
G18

Keywords:
Credit cards
Payments
Two-sided markets

ABSTRACT

We build a model of credit card pricing that explicitly takes into account credit functionality. In the model 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. 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 payment 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 payment like cash and debit cards which they believe to be more efficient. The usual suspects are the credit card interchange fees set by MasterCard and Visa, the transfer fees paid by the banks of the retailers (acquirers) to the banks of the cardholders (issuers), and which are often considerably higher than those for debit cards.[1] In the past several years, there have been more than 50 lawsuits concerning interchange fees filed by merchants and merchant associations against card networks in the United States, while in about 20 countries public authorities have take 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 and regulation. The point of departure from the existing literature on interchange fees 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).[2] The models in this literature have essentially focused on the choice between payment cards (which could just as well be debit cards) and cash. 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.

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 could capture a range of different types of purchases (such as unplanned purchases, impulse purchases and large purchases) for which the consumer does

☆ We benefited from the helpful comments of an anonymous referee, 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). Although both authors have consulted for Visa, this article was not funded in any way by Visa and the views expressed here are strictly those of the authors.

* Corresponding author. Tel.: +65 6516 3949; fax: +65 6775 2646.
E-mail addresses: rochet@cict.fr (J.-C. Rochet), jwright@nus.edu.sg (J. Wright).

[1] Unregulated credit card interchange fees are typically between 1% and 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).

[2] See Baxter (1983) for a much earlier treatment, and Rochet (2003) for a unified treatment.

0378-4266/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2010.02.026

not have the cash or funds immediately available to complete the purchase or for purchases for which the deferment of payment facilitates the transaction. Thus, offering credit allows an individual merchants to make sales that they otherwise would not make. The ability to make these incremental sales is, we think, the major reason explaining why merchants accept credit cards and indeed are willing to pay higher fees to do so compared to the fees paid to accept debit cards, and why prior to the widespread use of credit cards, store credit was much more widely used than today (Evans and Schmalensee, 2005, pp. 48–51).

For ordinary purchases, we assume credit cards are inefficient compared to pure payment technologies given we assume there are additional costs of transacting with credit cards.[3] As a result, in our model, 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" , 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 who 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-term) consumer surplus, our theory can provide a rationalization for placing a cap on interchange fees.[4] 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.

As we noted above, in the existing literature on interchange fees, the ability of merchants to make sales to consumers who would not buy if not for the availability of credit facilities is notably absent. Among the few related papers to explicitly model the extension of credit are Chakravorti and Emmons (2003), Chakravorti and To (2007), and Bolt and Chakravorti (2008). We view these studies as been complementary to ours. Rather than focusing on interchange fees and their regulation, the main focus in these papers is to provide a positive theory of the pricing and usage of credit cards. In particular, they are more explicit in modelling the different roles of credit, whereas we simply assume credit allows additional purchases to be made a fraction of the time. On the other hand, they do not consider store credit. The existence of store credit in our framework means credit cards will not have any effect on aggregate consumption.

For instance, Chakravorti and To show how the ability of a merchant to accept credit cards of liquidity constrained consumers leads to inter-temporal business stealing, in which merchants that attract additional sales today by accepting credit cards do so at the expense of rival businesses whom might have attracted the same sales tomorrow. This effect complements the intra-temporal business stealing effect that is present in our paper (and in earlier papers, starting with Rochet and Tirole, 2002). On the other hand inter-temporal business stealing is not considered in our setting. The existence of store credit means liquidity constrained consumers do not need to delay their purchase if a merchant does not accept their credit card.

Another difference between our setting and these other models of credit cards is that to simplify the model we do not incorporate the interest paid by consumers on their credit balances into our model, which would introduce another price to be determined (i.e. the interest rate on these balances). According to Chakravorti and Emmons, in the US, over 75% of US card issuer revenue is derived from cash-constrained consumers. They explore the interesting possibility that issuers would want to entice convenience users with low cardholder fees in order that some of these users would end up paying high interest on credit card debt (as revolvers). By ignoring interest income, we assume interest income is not relevant for the setting of interchange fees. This assumption is appropriate if issuing is competitive, so that there is no excess profit from credit card debt, and consumers can easily transfer their balances across issuers to obtain the best interest rate. The assumption is also appropriate to the extent consumers are rational and take into account any excess interest charges in their initial choice of issuer so that an issuer cannot benefit by shifting its cardholder fees to interest charges. In any case, in our model, interchange fees are set by card network operators (e.g. MasterCard and Visa) to maximize the volume of their credit card transactions, so these payment networks are not motivated by interest income as they would be in a three-party scheme (or an issuer-controlled system).

Closer to our own work is a recent paper by Bolt and Chakravorti who do endogenously solve for the optimal price structure of a card scheme, albeit of a three-party scheme, which sets prices to cardholders and merchants directly.[5] In particular, they model the insurance role of credit cards and how this affects the pricing structure of credit cards. In their model, consumers participate in credit cards to insure themselves against three types of shocks – income shocks, theft and merchant match, which can affect the possibility of consumption. In our setting, the availability of store credit would again be the relevant alternative, rather than a loss of consumption.

The rest of the paper proceeds as follows. Section 2 presents our benchmark theoretical model in which the proportion of consumers holding credit cards is taken as given. In Section 3 we determine the equilibrium level of the interchange fee, and develop a regulatory cap on interchange fee that raises (short-term) consumer surplus. Section 4 considers an extension of our benchmark model to endogenize the fraction of consumers that hold credit

[3] Estimates of the cost of using different forms of payment are discussed in Humphrey (2010).

[4] 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 consumer 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.

[5] Their analysis of price structure therefore does not correspond to our analysis of interchange fees (which are assumed to maximize the volume of credit card transactions) if bank margins in their setting vary with cardholder fees. Rather, assuming acquiring is perfectly competitive (as we do), their approach would correspond to the setting of interchange fees by an issuer-controlled card system.