EUROPEAN CENTRAL BANK
EUROSYSTEM

WORKING PAPER SERIES
NO 1682 / JUNE 2014

PAYING FOR PAYMENTS

FREE PAYMENTS AND OPTIMAL INTERCHANGE FEES

Søren Korsgaard

In 2014 all ECB publications feature a motif taken from the €20 banknote.

RETAIL PAYMENTS AT A CROSSROADS:
ECONOMICS, STRATEGIES AND FUTURE POLICIES

Retail payments at a crossroads: economics, strategies and future policies

This paper was submitted and accepted for the bi-annual retail payments conference titled "Retail payments at a crossroads: economics, strategies and future policies" organised by the European Central Bank (ECB) and the Banque de France (BdF), on 21 and 22 October 2013 in Paris. The aim of the conference was to bring together academics, regulators and market participants to discuss possible developments and dynamics that will shape the future retail payments landscape.

Retail payments provide a very important, yet not the most visible, infrastructure for the operation of the real economy. The way economic actors pay is not only important from a theoretical point of view but also from a policy perspective as the costs for providing payment services are substantial in most countries. As we are witnessing a transformation in these markets from traditional paper-based payment instruments towards electronic means of payments, the interaction between market forces and regulatory initiatives continues to be a determining factor for the future. This interaction, and the possible policy conclusions stemming from it, was the main theme of this conference.

The selection and refereeing process of this paper has been carried out by the conference organisational team composed of experts from both organising institutions. As the conference was organised with a focus on the key issues above, papers were selected not only based on their standalone quality but also considering the relevance of the research subject to the themes of the event. Following the conference the selected papers have been asked to be revised according to the discussants assigned to the respective paper.

The paper is released in order to make the working papers and accompanying research submitted to the conference publicly available. All working papers submitted to the conference can be found at
http://www.ecb.europa.eu/events/conferences/html/131021_ecb_bdf.en.html.

Acknowledgements

The author would like thank Thomas Gehrig (discussant) and participants at the ECB-Banque de France Retail Payments Conference (21-22 Oct 2013), colleagues and seminar participants at Danmarks Nationalbank, and Lasse Heje Pedersen for helpful comments.

Søren Korsgaard

Copenhagen Business School and Danmarks Nationalbank; e-mail: sko.fi@cbs.dk

Address Kaiserstrasse 29, 60311 Frankfurt am Main, Germany
Postal address Postfach 16 03 19, 60066 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Internet http://www.ecb.europa.eu

ISSN 1725-2806 (online)
ISBN 978-92-899-1090-3
EU Catalogue No QB-AR-14-056-EN-N (online)

Abstract

Do consumers and merchants use the most efficient payment instruments? I examine how *interchange fees*, which are fees paid from merchants' banks to consumers' banks when card transactions take place, influence the choice between cash and payment cards. I show that when consumers do not pay transaction fees to banks - a common feature in bank contracts - card use is declining in interchange fees, and surcharging does not neutralize interchange fees. According to my model, banks set interchange fees at too high a level, resulting in too few card payments. I derive an optimal interchange fee which depends only on the relative costs of producing cash and card payments and can be used by regulators to assess privately set interchange fees. When calibrated to cost data, the model implies an optimal fee that is low and may even be negative. The findings are consistent with empirical evidence of high card usage in countries with no interchange fees and have implications for the regulation of interchange fees.

Keywords: Financial regulation; Interchange fees; Payments

JEL-classification: E42; G21; G28

Non-technical summary

When consumers make a card payments, an *interchange fees* is typically paid from merchants' banks to the consumers' banks. Interchange fees can, in principle, help remedy an imbalances in the distribution of cost and benefits across the consumer's and the merchant's sides of the market. As an example, suppose that consumers derive relatively low benefits from the use of payment cards relative to cash, whereas merchants enjoy substantial cost savings from accepting cards. If, then, the bulk of the cost of producing card payments is borne by consumers' banks, it may not worthwhile for consumers' banks to produce card payments even though the joint benefits of consumers and merchants exceed the joint costs of the banks. An interchange fee is one means of solving this problem.

Much of the earlier literature on interchange fees highlights this potential efficiency. It is argued that interchange fees, if set optimally, increase the use of payment cards. A number of papers also suggests that even though interchange fees are often agreed upon collectively (multilaterally) by banks, they do not function as monopoly prices. This feature is due to the two-sided nature of the market for payment instruments. Furthermore, it is not clear that banks systematically set interchange fees either too high or too low relative to the social optimum. Moreover, even if private interchange fees were biased, it would not be possible to a regulator to set an optimal fee because such a fee would depend on unobservable factors. Taken together, these conclusions make a weak case for regulation of interchange fees.

In this paper, I show that these conclusions are reversed when banks do not charge transaction fees to consumers. A central assumption in the earlier literature is that higher interchange fees translate into lower marginal costs for consumers when paying by card. In practice, however, payment services are often provided either for free or against periodical fees: Consumers' marginal cost is therefore zero, independent of interchange fees. The implication is that as interchange fees increase, merchants' costs of accepting card payments rise, and fewer merchants will accept cards, whereas consumer behavior will remain unchanged. As a result, higher interchange fees are associated with less card usage. I also find that banks systematically set interchange fees above the social optimum. Interchange fees do, in fact, function as monopoly prices. The upshot is that there will be fewer card payments than is socially optimal.

A striking result of the model is that the socially optimal fee depends only on the difference in marginal cost of producing card and cash payments. Given adequate data on the cost of producing payments, it can therefore be calculated. The paper therefore not only shows that interchange fees are biased, but provides a potential method for determining the size of the bias. These findings suggest a positive role for the regulation, in particular the lowering of interchange fees as has been proposed by a number of authorities.

I also analyze the effects of surcharging. In my model, surcharging does not "neutralize" the effect of interchange fees, as is often argued in the related literature. Intuitively, this is because interchange fees are not passed on to consumers in the first place. In the model, surcharging has two effects. First, surcharging makes it less attractive for consumers to use cards. Second, banks will optimally choose a lower interchange fee, making it more attractive for merchants to accept cards. On balance, though, surcharging will be associated with lower card usage.

Finally, I calibrate the model to data on the cost of producing payments. Analysis of the calibrated model shows that optimal fees are likely to be close to zero, possibly even negative. This is because the optimal interchange fee depends on the difference between the marginal cost of producing card payments and the marginal cost of producing cash payments. Since the marginal cost of producing card payments is lower than that of cash payments, a social planner will therefore want to support to use of the use of cards relative to cash. As the model predicts card usage to be a decreasing function of the interchange fee, the optimal response is to set a low or even negative fee.

1 Introduction

A key role of banks is to provide payment services. The cost of producing these services is non-negligible. Based on a joint microeconomic study conducted by central banks in 13 European countries, the European Central Bank (Schmiedel et al., 2012) estimates the real cost of producing payments to be almost 1.0 per cent of GDP on average. More than 70% of this cost is related to the production of payments by cash or payment cards, and about half of the cost is incurred by banks. In spite of these costs, bank contracts often provide little incentive for consumers to choose payment instruments that reflect their cost of production. Bank contracts commonly involve periodical fees which give access to the use of basic payment products such as debit cards, but no transaction fees, implying that consumers' marginal cost of paying is zero. Borzekowski et al. (2008) cite evidence that just 15% of US banks charge consumers for debit card transactions.[1] A similar pattern is found in many European countries.[2]"

The literature on payments points to a mechanism known as an *interchange fee* that can potentially direct consumers and merchants towards the use of the most efficient payment instruments. An interchange fee is a transfer from a merchant's bank to a consumer's bank when a card transaction takes place. A simple example may illustrate why such a fee can improve welfare. Suppose a consumer values the ability to pay by card at 0.5, while a merchant enjoys savings of 1.0 when payment is made by card. If the consumer bank's cost is 0.75, and the merchant bank's cost is 0.25, it is beneficial for the transaction to take place. In the absence of a transfer from merchants or their banks to consumers or their banks, however, a card transaction will not be feasible, since the consumer's bank cannot cover its costs by charging the consumer. The presence of an interchange fee could in principle solve this problem (as could a card payment rebate offered by merchants). Banks and card networks use such an argument to persuade regulators and competition authorities that interchange fees are necessary to induce consumers to use cards. Merchants, on the other hand, view interchange fees as a collusive device (Wright, 2012). Merchants can ill afford to refuse card payments as it would entail a loss of business, and banks exploit this by setting high interchange fees. As interchange fee levels are typically agreed upon collectively by banks, they determine a lower bound for merchants' cost of receiving payment by card. The issue has likewise attracted the attention of regulators, who have intervened by regulating fees. In the US, the Durbin Amendment to the Dodd-Frank Act introduced caps on interchange fees, with the Federal Reserve being assigned responsibility for setting standards for interchange fees. In Europe, the European Commission has required Visa and MasterCard to lower fees for cross-border transactions, and has in 2013 proposed a regulation of interchange fees that would cap fees at just 0.2% in the case of debit cards and 0.3% in the case of credit cards.

The debate on the role of interchange fees raises a number of questions. For instance, what is the effect of interchange fees on cash and card usage? What would be a socially optimal fee, and what information would a regulator need to have to calculate it? Are privately set fees biased relative to the social optimum? I arrive at strong answers to these questions. I show that when consumers do not face transaction fees, card usage is declining in interchange fees. Banks will set interchange fee which exceeds or equals the socially optimal fee. I further derive a socially optimal fee which depends only on the relative costs of producing cash and

[1] The number is based on data from 2004 and covers only PIN-based debit card transactions from which banks, according to the authors, derive less interchange fee revenue than for signature-based transactions. Banks would therefore seem to have an incentive to charge consumers fees for exactly these transactions.

[2] As an example, Jonker (2013) describes the typical fee structure in the Netherlands as follows: "[...] banks call these fees payment package fees. The fees are linked to a payment package including the use of a current account, access to internet banking and a debit card which consumers can use for making debit card payments or to use ATMs. Credit cards are provided optionally at extra cost. Consumers do not pay any transaction fees for card payments, cash withdrawals or cash depositions at the ATM [...] [banks] tend not to make consumers directly aware of the costs associated with their payment behaviour or to provide any incentives towards more cost efficient payment behaviour.

card payments. Taken together, these results have important implications. Since higher fees are associated with less card usage and banks set too high a fee, there will be too few card payments in the economy if interchange fees are not regulated. Moreover, interchange fees can be regulated, and the existence of a cost-based optimal fee provides a means for evaluating fees. I also consider the role of alternatives to interchange fee regulation, for instance permitting merchants pass on, or *surcharge*, fees to consumers. The welfare effects of permitting surcharging are ambiguous. When fees are unregulated, permitting merchants to surcharge will lead banks to lower their fees, which is positive from a welfare perspective. On the other hand, the surcharge itself causes some consumers to pay using cash, even though they prefer card payments. Moreover, it is no longer possible to identify a socially optimal fee when surcharging is permitted. These findings contrast with those in the earlier literature. The more traditional view is that interchange fees foster card adoption and so lead to increased card usage. In that view is not clear that privately set interchange fees are systematically biased relative to the social optimum and, in any case, the socially optimal fee will depend quantities, which are unobservable to regulators.

The model is based on an analytical framework similar to that in Rochet and Tirole (2002, 2011) and Wright (2004, 2012), who extend the original analysis of interchange fees due to Baxter (1983). It involves three types of agents: consumers, merchants and banks. Consumers and merchants derive some benefit from using cards instead of cash. These benefits are heterogeneous; some consumers and merchants have a preference for card payments whereas others may prefer cash. Consumers choose which merchants to frequent based on their prices and whether they accept cards. Merchants decide whether to accept payment cards based on the benefits they enjoy, the fees they must pay, and the fraction of consumers preferring to pay by card. Banks face costs related to the production of card and cash payments, and set interchange fees in order to maximize their joint profits, which will depend on the fraction of consumers and merchants respectively using and accepting cards.

The critical assumption in this paper is all consumers have payment cards and do not face transaction costs when paying. In the reference literature, consumers' banks are assumed to charge consumer a transaction fee equaling their cost of producing payment services, plus a markup reflecting imperfect competition. Higher interchange fees therefore translate directly into lower consumer transaction fees and a greater willingness to use cards among consumers. If merchants are homogeneous (Rochet and Tirole, 2002), the effect is that card usage is increasing in interchange fees, at least up to a point where merchants refuse to to accept cards. If merchant heterogeneity is incorporated (Wright, 2004), fewer merchants will accept cards as interchange fees increase, and there will no longer be a monotone relationship between interchange fees and card usage. When consumers do not pay transaction fees, the first of these mechanisms breaks down: Lower interchange fee no longer induce consumers to pay by card instead of using cash. The only effect of interchange fees is to limit merchant acceptance, which means that card usage decreases in interchange fees.

A similar reasoning underlies the result that the privately set interchange fee exceeds the socially optimal fee. The market for payment instrument is an example of a *two-sided* market (Rysman, 2009; Rochet and Tirole, 2006); the total benefits enjoyed by consumers depends on the fraction of merchants accepting cards, and the total benefits enjoyed by merchants depends on the fraction of consumers using cards. It is this feature of the market which - when banks do charge transaction fees - makes interchange fees different from prices in one-sided markets. Even if they are set collectively by banks, the outcome is not monopoly pricing (Schmalensee, 2002; Wright, 2004), and one cannot conclude that privately set interchange fees will necessarily be biased relative to the social optimum.When banks do not charge transaction fees, however, this logic fails. Since consumers' payment decision is unaffected by the level of interchange fee, the intuition is that banks will effectively act as monopolists towards merchants. The upshot is that banks set interchange fees too high, a result echoed (albeit for other reasons and with different implications) in Rochet and Tirole (2002).

The empirical evidence is consistent with the findings. While data on interchange fee levels is scarce, differences in regulatory regimes across countries provide an indirect test. Börestam and Schmiedel (2011) report that in 2006, four European countries had national card schemes, which operated entirely without interchange fees, namely those in Denmark, Finland, Luxembourg and the Netherlands. If it were true that interchange fees were critical to the adoption of payment cards, as card networks and banks argue, one would expect only weak card adoption in these countries. On the other hand, if it is simultaneously true that card usage is decreasing in fees and that privately set fees exceed those set by a regulator, then one might expect regulation to have a positive impact on card usage. Chart 1 shows card usage in the countries that were part of the euro in 2006, plus Denmark. The countries which have historically operated payment card schemes without interchange fees also happen to be those in which card usage is greatest.

[Figure 1 about here.]

Still, a common problem with models of interchange fees is that even if one can show the privately set fee to be biased, the socially optimal fee will generally depend on unobservable quantities such as the distribution of consumer benefits. This renders it of limited value to a regulator. One test that has proposed is the so-called merchant indifference test, or tourist test, whose merits is explored by Rochet and Tirole (2011). The merchant indifference test suggests that fees should be set such that merchants are indifferent between receiving cash and card payments. The test motivated by complaints that merchants face excessive card fees (e.g. Vickers, 2005), sometimes referred to as the *must-take-cards*-problem. Informally, the argument is that merchants cannot refuse to accept cards if enough consumers wish to pay using cards, hence allowing banks to demand excessive fees (Wright, 2012). In my baseline model, I derive an optimal interchange fee that is function only of the costs of producing payments, and am able to compute an optimal interchange fee based on data from Danmarks Nationalbank (2011), which is part of the ECB cost of payments study cited above. The resulting optimal interchange fee is negative, though close to zero. It is negative because banks underprice cash payment services and therefore essentially subsidize cash usage. The optimal response for social planner is require a low cost of card payments. I also compare the obtained interchange fee level to the fee level derived from both a method akin to that used by industry and the level based on merchant indifference test.

The existence of strictly cost-based optimal fee hinges on a number of model assumptions, and can no longer be obtained if these are altered. For instance, when merchants are assumed to have fixed costs associated with receiving card payments, the optimal fee will no longer be purely cost-based. In that particular case, the optimal fee to equal the cost-based fee plus another term, which I show to be always negative. Thus, the cost-based fee is still useful as it provides an upper bound for the optimal fee.

The conclusions of the model are altered somewhat in the special case of homogeneous merchants. In that case there is a whole interval of socially optimal fees, and among these banks will choose the highest fee. The effects of interchange fee regulation are then purely distributional, a matter of whether merchants or banks pay for payments. In fact, much of the debate on the effects of interchange regulation has emphasized the distributional consequences (Wang, 2012).

In order to obtain quantitative insights, I calibrate the model to data on the cost of producing payments obtained from the Danish part of the European Central Bank (Schmiedel et al., 2012). This results in an optimal interchange fee that is negative. I compare this to fees obtained by other methods. Interestingly, the fee implied by the tourist test is about 0.2%, which is the same figure as that agreed to by the card companies and the European Commission for debit card transactions and suggested as an upper bound for permitted in interchange fees in a proposal for regulation issued by the European Commission. The model predicts a privately optimal fee of about 1.3%.

I further examine a number of extensions to the model, most importantly what happens when merchants are permitted to surcharge fees. Surcharging is interesting because it has the potential to neutralize the effect of interchange fees (Gans and King, 2003; Zenger, 2011). Interchange fee regulation becomes more complicated when surcharging is permitted for at least two reasons. First, it is not straightforward to establish the conditions under which merchant surcharge, and there may multiple equilibria when surcharging is permitted. Empirical studies confirm that surcharging, where permitted, is not uniformly practiced (Bolt et al., 2010; Jonker, 2011). Second, if merchants do surcharge, the optimal interchange fee is no longer the same as in the no-surcharging case, and it is no longer possible to derive a cost-based optimal interchange fee. The analysis of the calibrated model shows the welfare effects of surcharging to be ambiguous. For positive merchant fees, welfare is greatest when merchants do not surcharge. The model shows the welfare effects of surcharging to be ambiguous. At low, but positive merchant fees welfare is greatest when merchants do not surcharge, and the model predicts that merchants will choose to surcharge at such fees. On the other hand, surcharging may improve welfare if merchant or interchange fees are unregulated. This is because, while welfare is lower for a given merchant fee, the presence of surcharging merchants will also induce banks to set a lower merchant fee than if merchants did not surcharge. On balance, the outcome might be an improvement in welfare.

As for the institutional setting, the model is best thought of as describing the market for debit card payments since debit card payments are more direct substitutes for cash payments than are credit card payments. The determination of interchange fees for credit card is slightly different (Rochet and Wright, 2010) as credit card payments involves both a payment component and credit provision component. Moreover, the assumption that consumers do not pay transaction fees is more appropriate in the case of debit cards, as credit card contracts typically do come with different kinds of consumer benefits. As for price structure, it is commonly observed that interchange fees are proportional fees rather than per transaction fees. This may in part reflect the underlying cost structure of providing the card service, for instance insuring consumers and merchants against card abuse, but may also reflect issues of market power (e.g. Shy and Wang, 2011). The issue of price structure is sidestepped in this paper via the assumption that all consumers face unit demand for a single good. Finally, the model assumes a linear city model of merchant competition as in Rochet and Tirole (2002). Other modes of competition are considered in Wright (2003), while competition between payment networks are discussed by Guthrie and Wright (2007); Chakravorti and Roson (2006). In the model, the effect of assumptions about competition is to provide a specific threshold: Merchants whose benefits exceed the threshold accept cards, and others do not. Under slightly the different assumption the same threshold can be shown to hold under other types of competition, including perfect competition (Rochet and Tirole, 2011), so the results of the paper are not a peculiar feature of such assumptions. Further references to the literature on interchange fees can be found in the surveys by Chakravorti (2003); Verdier (2011). Börestam and Schmiedel (2011) provide an overview of institutional details in the market for payments cards.

The rest of the paper is organization as follows. The baseline model is presented in section 2. Results are given in section 3, and extensions to the model are discussed in section 4. A policy analysis in a calibrated version of the model is conducted in section 5, which also presents evidence of high card usage in countries without interchange fees. Section 6 concludes. All proofs are provided in the appendix.

2 Model

The model features three types of agents: consumers, merchants and banks. Merchants produce goods, which are consumed by consumers. Consumers can choose to pay for these using cash or cards. Banks collectively set merchant fees, and consumers and merchants respectively decide whether to use payment cards and accept card payments.

There is a unit length continuum of pairs of merchants. For each merchant pair, the two merchants (or sellers, "S") enjoy identical benefits b_S from accepting cards. Two merchants are located at opposite ends of a line segment (linear city) of unit length, and compete for consumers who are uniformly distributed across the segment. Merchants produce a good, which costs γ to produce and set prices p_i $i = 1, 2$ to maximize profits. In addition, merchants choose whether to accept card payments. If they accept payment by card, they must pay a merchant fee of m when consumers use cards. In addition, merchants may face a fixed cost K if they choose to accept cards. Merchants are heterogeneous across pairs, with benefits distributed according to a cumulative distribution function $G(b_S)$ with full support on the interval $[b_{S,min}, b_{S,max}]$. The benefits can be thought of as a merchant's net savings from receiving payment by card rather than cash. It is assumed that merchants are prohibited from surcharging. This assumption is relaxed in a later extension to the model.

For each merchant pair, the economy is inhabited by a unit mass of consumers (or buyers, "B"). Each consumer faces unit demand for goods produced by a merchant, and is located between the pair of competing merchants. Consumers are heterogeneous in the benefits they derive from paying by card, and all have a payment card. The assumption throughout most of the analysis will be that consumers do not face adoption costs. In the first part of the results section, however, I address the question of existence of equilibria in the more general case where consumer adoption costs are present. The consumer benefits, which can be interpreted as consumers' willingness-to-pay for paying by card, are denoted b_B and are distributed according to cumulative distribution function $H(b_B)$ with full support on the interval $[b_{B,min}, b_{B,max}]$. Consumers face a distance cost of t per unit of distance located from a merchant.

The cumulative distribution functions of the consumer and the merchant are assumed to be twice continuously differentiable. For illustrative purposes, results are sometimes shown under the assumption that merchant and consumer benefits are uniformly distributed. A sufficient condition for guaranteeing concavity of the maximization problems is that the hazard rate $\frac{g(b_S)}{1-G(b_S)}$ be increasing.[3]

Banks produce payment services, and set merchant fees collectively in order to maximize joint profits. No distinction is made between consumers' and merchants' banks. This can be motivated e.g. by assuming that merchants' banks are perfectly competitive. In that case the merchant fee can be written as the sum of the interchange fee and the merchant banks' marginal cost of producing card payments; distinguishing between consumer and merchant banks is thus of no economic consequence. To keep notation simple, I will therefore treat consumers' and merchants' banks as one with the implicit understanding that there is a one-to-one relationship between the interchange fee and the merchant fee. Banks face variable costs of c_C, which are proportional to the fraction of card payments in the economy. Banks obtain card revenue from merchants, and this revenue is equal to the product of the merchant fee m and the fraction of card payments in the economy. In addition, banks face costs related to the production of cash payment services, which are proportional to the fraction of cash payments in the economy with transaction costs c_D.[4]

Banks' joint objective is therefore to maximize

$$\Pi = (1 - H(b_B^\star))(1 - G(b_S^\star))(m - c_C) - (1 - (1 - H(b_B^\star))(1 - G(b_S^\star))c_D \qquad (1)$$

where $b_B^\star$ and $b_S^\star$ denote the thresholds above which consumers and merchants use and accept cards. For much of the paper the assumption is that $b_b^\star = 0$, as consumers do not pay for the use of payment instruments, though this assumption must be altered e.g. when consumers are subject to surcharging. To simplify notation going forward, I will sometimes refer to the total fraction of card payments as

[3]Such a condition is generally applied in the reference literature. In fact, this condition is more than sufficient as weaker conditions can be imposed. I discuss these in the relevant proofs.

[4]In the model I assume for simplicity that the social costs of producing payments, c_C and c_D, are the same as banks' private costs. More realistically, banks will have some cost recovery from consumers, e.g. in the form of periodical fees. In the numerical analysis of the model I will take into account the difference in private and social cost.

$$\mu(m) \equiv (1 - H(b_B^\star))(1 - G(b_S^\star)) \tag{2}$$

The presence of cash is taken for granted, that is, banks cannot choose not to provide cash payment services. The individual rationality constraint facing banks is therefore that they must be no worse off providing card payment service than they would be in a pure cash economy, i.e.

$$\mu(m)(m - c_C) - c_D(1 - \mu(m)) \geq -c_D \tag{3}$$

The social planner is assumed to maximize the total benefits resulting resulting from the use of payment cards less the costs associated with the production of payment. The social planner's objective function is:

$$\begin{aligned} W \quad = \quad & (1 - G(b_S^\star)) \int_{b_B^\star}^{b_{B,max}} b_B dH(b_B) + (1 - H(b_S^\star)) \int_{b_S^\star}^{b_{S,max}} b_S dG(b_S) \\ & -\mu(m)c_C - (1 - \mu(m))c_D - (1 - G(m - \delta)]K \end{aligned} \tag{4}$$

In derivations, I will assume that that $b_{B,min} < 0 < b_{B,max}$ and that $b_{S,min} < b_S^\star < b_{S,max}$ in order to avoid having to deal with "corner solutions" in which either all merchants or consumers use cards.[5]

3 Results

I first examine the question of card adoption in the economy. In particular, when is an equilibrium guaranteed to exist in the most general case where both consumers and merchants face adoption costs? This represent a departure from the assumption made in the baseline model described above (where consumers adoption costs are not present) and in the related literature where neither merchants nor consumers face any adoption costs.[6] The introduction of consumer adoption costs adds extra generality - it makes the market truly two-sided in the sense that consumers' adoption decision will then depend on the fraction of merchants accepting cards, while merchants' acceptance decision depends on the fraction of consumers with cards - but comes at the expense of weaker results. It turns out that one can specify the conditions under which an equilibrium exists, but not even a unique equilibrium can be guaranteed. For the remainder of the analysis I therefore dispense with the assumption that consumers face adoption costs, but mostly retain the assumption of merchant adoption costs. This simplification can be justified by the argument that consumer adoption costs are likely to be negligible in practice. It is common for consumers to automatically be provided with a payment card as part of their basic banking services, and so the critical decision is the usage decision, which depends on the marginal cost of paying with different payment instruments, and not the adoption decision.

When consumers face fixed adoption costs, which I denote k, their decision to get a card will depend not only on the price of paying by card, which is zero by assumption, but also on the fraction of merchants accepting cards. Merchants decision, in turn, depends on the fraction on consumers using cards; this is the case independently of whether merchants face a fixed adoption cost. The effect of introducing consumer adoption costs is to make the market becomes two-sided such that each side of the market, consumers and merchants, makes a participation choice that depends directly on the participation choice of the other side.

[5]This issue will be taken into account when I examine the model numerically; in that case, it will become apparent that the last assumption is not necessarily satisfied for low values of m at which all merchants accept cards.

[6]An exception is McAndrews and Wang (2012), who employ a model of quite different style than is the standard in the literature.

To show the existence of equilibria it is necessary to analyze the participation choices of merchants and consumers. In the case of consumers this is straightforward. A consumer will adopt a payment card if

$$b_B(1 - G(b_S^\star)) \geq k \tag{5}$$

i.e. the total benefit exceeds the adoption cost. I let $b_B^\star$ denote the adoption threshold for consumers, with $b_B^\star = \frac{k}{1-G(b_S^\star)}$. Here $b_S^\star$ denotes the threshold above which merchants accept cards. For merchants the problem is slightly more involved. Merchants acceptance threshold is given in the following lemma.

Lemma 1. *A pair of merchants accepts payment by card provided that the following condition is satisfied*

$$b_S \geq b_S^\star = m - E[b_B|b_B \geq b_B^\star] + \frac{3t(1-\sqrt{1-\frac{2K}{t}})}{1-H(b_B^\star)} \tag{6}$$

This condition has several implications. If there are no fixed costs, i.e. $K = 0$, the condition reduces to

$$b_S \geq m - E[b_B|b_B \geq b_B^\star] \tag{7}$$

which is the same acceptance condition as in Rochet and Tirole (2002); and while the derivation in the appendix relies on the linear city model of competition among merchants, the same condition can be shown to hold under perfect competition or local monopoly (Rochet and Tirole, 2011). The condition implies that merchants might accept cards even if it increases their costs. If there is no fixed cost, though, the result has a somewhat counterintuitive implication: As the card adoption threshold for consumers increases and fewer consumers use cards, more merchants will accept cards. This conclusion no longer follows automatically with a fixed cost, since the cost must be spread among fewer card users. While $E[b_B \mid b_B \geq b_B^\star]$ is increasing in $b_B^\star$, $1 - H(b_B^\star)$ is decreasing with opposite effect on card acceptance. Differentiation of the expression with respect to t shows the expression to be declining in t, meaning that more merchants will accept cards. In fact, the acceptance threshold is only defined for $K \leq 1/(2t)$. This makes sense since for any higher K, merchant profits would be negative if merchants accepted cards.

An equilibrium requires that the fractions of consumers and merchants desiring to adopt and accept cards must agree with each other as the fraction of consumers adopting cards depend on the fraction of merchants accepting cards and vice versa. In equilibrium, if a a certain fraction of consumers find it optimal to adopt cards given the fraction of merchants accepting cards, it must be the case that exactly that fraction of merchants do, in fact, accept cards given the choice of the consumers. This can be formulated as a fixed-point problem, here viewed from the perspective of the merchants. There is, of course, a trivial fixed point, namely that in which cards are neither adopted by consumers nor accepted by merchants. In looking for non-trivial fixed points, I therefore focus on benefit intervals of the form $S = [b_{S,min}, b_{S,max} - \epsilon]$, where ϵ is a positive number. One can then formulate the problem as one of finding a fixed point of the function

$$f(b_S) = m - E\left[b_B \mid b_B \geq \frac{k}{1-G(b_S)}\right] + \frac{3t\left(1-\sqrt{1-\frac{2K}{t}}\right)}{1-H\left(\frac{k}{1-G(b_S)}\right)} \tag{8}$$

Lemma 2. *A fixed point exists whenever for all b_S in S:*

$$b_{S,min} \leq m - E\left[b_B \mid b_B \geq \frac{k}{1-G(b_S)}\right] + \frac{3t\left(1-\sqrt{1-\frac{2K}{t}}\right)}{1-H\left(\frac{k}{1-G(b_S)}\right)} \leq b_{S,max} - \epsilon \tag{9}$$

The condition is a sufficient condition, which guarantees the existence of an equilibrium. An equilibrium can certainly exist in cases where the condition is not satisfied. For instance, it is possible for an equilibrium to exist in which $f(b_S) \leq b_{S,min}$. That would correspond to a situation in which all merchants accept cards, and consumers for whom $b_B \geq k$ use cards.

More generally, the expression for $f(b_S)$ highlights that two economic effects are present as b_S is varied. If b_S decreases, i.e. more merchants are willing to accept cards, more consumers will also be induced to get a card. The first effect is then that the average consumer who prefers cards to cards has a weaker preference for cards than before, since the marginal card users are those who do not have a strong preference for cards. This means that fewer merchants will want to accept cards. The second effect is that as more consumers wish to pay by cards, merchants can spread their fixed cost across more consumers, making them more likely to accept cards. These offsetting effects suggest that there might be two equilibria, one in which cards are widely used and one which they are less widely used. In section 5.4, I show an example of this using plausible parameter values. If one were to think of the model in dynamic terms, only the high card usage equilibrium would be a stable equilibrium.

It is also possible to say something about the effect of changing parameter values. An increase in k, for instance, has the immediate effect of reducing the fraction of consumers using cards. Following the same logic as before this has two effects: Only those consumers with the strongest preferences cards remain, making it more attractive for merchants to accept cards, but the fixed cost element increases. Indeed, there is a limit to how high k can become, as the fixed cost term will tend to infinity as k increases and b_B approaches $b_{B,max}$. For the same reason, ϵ cannot be chosen arbitrarily small, since $f(b_S)$ tends to infinity for b_S too close to $b_{S,max}$. In plain terms, if cards are to adopted it must not to expensive for consumers to obtain a card. It is likewise clear that merchant fixed costs likewise cannot be too high, while conversely it is more likely that equilibria are guaranteed to exist when the maximum benefits enjoyed by consumers or merchants are high.

It is instructive to consider what happens when b_S is close to $b_{S,min}$, meaning that virtually all merchants would like to accept cards. Both m and the fixed cost term are positive. Presumably, some merchants dislike cash, so the only way for for $f(b_S)$ to be less than $b_{S,min}$ is if the average benefit to cardholders is very high. This occurs when k is large, since in that case only individuals with a strong preference for card payments adopt cards. Hence, for large values of k an equilibrium may not exist. Conversely, $f(b_S)$ may exceed the upper bound. This happens if either merchant fees or fixed costs are too high, i.e. if accepting cards becomes too expensive for merchants. Now, consider what happens if the adoption threshold for merchants increases. This has two effects of opposite sign. First, as fewer merchants accept cards, it becomes less attractive for consumers to use cards. Only those with a strong preference for cards remain, leading to an increase in the average cardholder benefit. As a result, f decreases. Second, as fewer consumers use cards, the fixed cost element increases in importance, as the cost is spread over fewer consumers. This leads to an increase in the value of f.

For the remainder of the analysis it will be assumed that fixed consumer adoption costs k are zero, which leads to stronger and more interesting results. As noted earlier, this assumption does not seem too unreasonable as a payment card is often automatically provided as part of a package of basic banking services. When $k = 0$ there is a unique equilibrium, which can be calculated. One can also make a clear statement about the effect of interchange fees on card usage, as detailed in the following proposition.

Proposition 1 (Card usage and interchange fees)**.** *When consumers do not face adoption costs, i.e.* $k = 0$*, card usage is declining in merchant fees.*

The intuition underlying this result is straightforward. When consumers' marginal cost of paying is zero, the interchange fee does not affect their incentive to pay by card. Those who prefer to pay by card will do so when they can, independent of the interchange fee. On the other hand, fewer merchants accept cards, and so a number of individuals who would have liked to pay by card are no longer able to. The result is fewer card payments.

This establishes that card usage is decreasing in merchant fees or, equivalently, in interchange fees, thus providing a possible explanation why countries without interchange fees also exhibit a high fraction of card payments. While the result does not have a direct bearing on the issue of whether privately set interchange fees are optimal, it does show that if privately set interchange fees exceed the social optimum, then the resulting outcome is too few card payments.

As an empirical matter, it is difficult to verify a link between interchange fees and card usage for at least two reasons. The first is absence of data; frequently, interchange fee levels are often not publicly disclosed. The second reason is that the observed interchange fees are endogenous. In a number of countries, however, fee levels have been imposed by regulation. This has often been the case in countries where national card schemes have been in existence, such as in the European countries discussed in the introduction. Empirical evidence on the link between interchange fees and card usage, while scarce, is consistent with the model's predictions. As described in the introduction, Börestam and Schmiedel (2011) report that in 2006, four European countries had national card schemes, which operated without interchange fees, namely Denmark, Finland, Luxembourg and the Netherlands. These were also among the countries with the most card payments. Other countries with relatively many card payments at the were France and Portugal. Börestam and Schmiedel (2011) also mention that interchange fees are regulated in France, and report fairly low interchange fee levels in Portugal compared to these in neighboring Spain where card payments are significantly rarer.[7]

A natural follow-up question is whether interchange fees as set by banks are biased relative to the social optimum? The answer is in the positive: Privately set fees exceed those set by the social planner, except in special cases. In order to arrive this conclusion, it is necessary to compare the optimal choices for the banks and the social planner. This is the content of the next lemma and proposition.

Lemma 3 (Banks' profit-maximizing merchant fee)**.** *Banks' profit-maximizing merchant fee is indirectly defined by the following condition:*

$$m = c_C - c_D + \frac{1 - G\left(m - E[b_B \mid b_B \geq 0] + \frac{3t\left(1-\sqrt{1-\frac{2K}{t}}\right)}{1-H(0)}\right)}{g\left(m - E[b_B \mid b_B \geq 0] + \frac{3t\left(1-\sqrt{1-\frac{2K}{t}}\right)}{1-H(0)}\right)} \tag{10}$$

The expression says that banks equate marginal revenue and costs. To help build intuition, it is interesting to consider how the expression looks under specific distribution assumptions. With uniformly distributed benefits, the expected benefits of card users are

$$E[b_B|b_B \geq 0] = \frac{1}{2}b_{B,max} \tag{11}$$

Ignoring merchant fixed costs, this means that the optimal privately set merchant fee can be shown to be

$$m = \frac{1}{2}\left(c_C - c_D + b_{S,max} + \frac{1}{2}b_{B,max}\right) \tag{12}$$

As will be shown in the following, neither terms involving consumer nor merchants benefits appear in the expression for the socially optimal fee. This is because banks are effectively acting as monopolists. If consumer prices for payment services are fixed, the market becomes one-sided, since an increase in merchant fees is not counterbalanced by a reduction in consumer

[7]Data on card usage across countries can be found in chart 3 in the cited paper. Interchange fee levels for Spain and Portugal are reported in charts 6 and 7.

marginal costs. Banks set a fee that exceeds their marginal cost, resulting in an underprovision of card payment services. The expression furthermore indicates that merchant fees are more responsive to changes in merchant benefits than to changes in consumer benefits. It also shows that merchant or interchange fees should be falling as the costs of producing card payments decline. However, privately set fees increase if consumers' or merchants preferences' change in favor of card payments. Such and argument could perhaps explain why merchant fees have not declined over time in the US in spite of falling technology costs (McAndrews and Wang, 2012).

Banks' optimal merchant fee can be compared to the corresponding first-order condition for the social planner, again ignoring merchant fixed costs.

Proposition 2. *When merchant fixed costs are zero, the socially optimal interchange equals the difference in banks' marginal cost of producing card and cash payments, i.e.*

$$m = c_C - c_D \tag{13}$$

The expression for socially optimal merchant profits is remarkably simple, and the optimal fee can be determined solely based on costs. At first glance, the optimization problem facing the social planner seems rather complicated and more so than that facing the banks, yet the resulting expressions are relatively similar, except for a single bias term. Intuitively, what the social planner is attempting is to put cards and cash on an equal footing by setting merchant fees to reflect the difference in marginal cost between the two. To the extent that card payments are cheaper to produce than than cash payments, merchants should be induced to accept cards via negative fees. This essentially the same as the standard socially optimal pricing strategy of setting price equal to marginal cost. Comparing the planner's problem with the banks also points to the sources of the bias in banks' fee. A key source, for instance, is that the total consumer benefits depend on the fraction of merchants accepting cards, and this loss to consumers from too few merchants accepting cards is not internalized by banks.

It is straightforward to verify that the banks' rationality constraint is satisfied with equality at the socially optimal fee. An immediate consequence of the preceding lemma and proposition is therefore that the profit-maximizing merchant fee always exceeds or equals (see the later discussion of the case of homogeneous merchants) the welfare-maximizing fee.

Proposition 3 (Profit-maximizing versus socially optimal merchant fees)**.** *Banks always set merchant fees that exceed or equal those set by a social planner.*

In the above I assumed no merchant fixed costs, i.e. $K = 0$, but an identical logic applies in the case of positive fixed costs for merchants. The formula for privately optimal interchange fee is unchanged, though the fee itself will change due to the effect of K on the merchant acceptance threshold. The formula changes, however, in the case for the socially optimal fee. It can be shown to be

$$m = c_C - c_D + \frac{K - 3t(1 - \sqrt{1 - \frac{2K}{t}})}{1 - H(0)}) \tag{14}$$

The resulting expression is no longer purely cost-based as $1 - H(0)$ depends on the distribution of consumer benefits. That in itself does not pose too rigorous an informational requirement for a social planner, since $1 - H(0)$ is just the fraction of card users in the economy, which might be observable. It is rather the presence of a variable such as t, which makes the expression unpractical. Aside from added simplicity, at least four arguments can be made for ignoring the fixed cost term, and using $m = c_C - c_D$ as the benchmark when evaluating merchant fees. First, K might be small in practice and therefore of little consequence. Second, including the fixed cost term implies an even lower socially optimal merchant fee. Third, at $m = c_C - c_D$ it might be the case that the merchant threshold is actually lower than $b_{S,min}$, in which case a further reduction in fees is of no consequence. Fourth, the first part of the expression merchant acceptance threshold, i.e. $m - E[b_B \mid b_B \geq 0]$, also holds under alternative assumptions about merchant competition (such as perfect competition), whereas the fixed cost component depends on the degree of competition, which may differs across industries and would therefore suggest a need for differentiated optimal fees.

To see that the inclusion of merchant fixed cost produces a lower socially optimal merchant fee, one needs to determine the sign of $K - 3t\left(1 - \sqrt{1 - \frac{2K}{t}}\right)$. This expression must be negative since

$$K - 3t\left(1 - \sqrt{1 - \frac{2K}{t}}\right) \leq 0 \Leftrightarrow 3t\sqrt{1 - \frac{2K}{t}} \leq 3t - K \tag{15}$$

Both sides of the last expression must be positive, if cards are accepted by merchants. In an equilibrium where merchants accept cards, it must be the case that $\frac{1}{2}t \geq K$, since otherwise merchants would be making negative profits. That means that $3t - K > 0$. Squaring both sides and simplifying the expression, results in the expression $-12t \leq K$, which is obviously true.

There is a special circumstance in which the privately set interchange fee is also socially optimal, namely when merchants are homogeneous, that is, they enjoy the same b_S across all merchant pairs.

Corollary 1. *When merchants are homogeneous, i.e. enjoy the same b_S across all merchant pairs, there is an interval of socially optimal interchange fees. Banks' choose the highest among these.*

An interpretation is that to the extent that merchants are relatively homogeneous, the exact level of the interchange fee may not matter too much from a welfare perspective, as long as it is within certain bounds. In that case the consequences of changing the interchange fee level are mainly distributional, and the real question is whether banks or merchants reap the benefits.

4 Extensions

In this section, I consider a number of extensions to the model. I first examine the consequences of permitting merchants to surcharge merchant fees. I next consider a variant of the model where banks must pay a fixed cost to operate the card payment system. On a related note, I then discuss the incentives for banks to invest in innovations that either enhance benefits for consumers and merchants or reduce costs. Finally, I digress slightly by offering a heuristic argument as to why banks' frequently seem to be making losses on payment services.

4.1 Surcharging

Where permitted and applied, surcharging has the potential to neutralise the effect of interchange fees. Two questions arise in the context of surcharging. First, will merchants surcharge if permitted? Second, what are the welfare consequences of surcharging? Both questions, it turns out, defy simple analytical answers, but the analysis of surcharging nevertheless provides valuable knowledge about the effects of surcharging. The numerical analysis in section 5.3 sheds further light on these questions.

In practice, surcharging comes in different varities. Some companies have been criticized for surcharging amounts that are seemingly unrelated to the merchant fees, they are paying.[8] In the following I look at the case where merchants are permitted to surcharge the amount of the merchant fee or not surcharge at all. Such a requirement is typically stipulated in surcharging regulation.

[8] In Europe, certain airlines have been criticized for such a practice.

Admitting the possibility of surcharging complicates the strategic options for merchants. Where before they faced a choice between accepting cards (and not surcharging) and not accepting cards, the presence of surcharging introduces a third option which is to accept cards and surcharge the amount of the merchant fee. To check whether or when a surcharging equilibrium exists, it is therefore necessary to compare merchant profits when both merchants surcharge with their profits when deviating by either not accepting cards or accepting cards and not surcharging. As in the no-surcharging case, this problem can be tackled by finding the values for b_S for which deviating to either not accepting cards at all or accepting cards and not surcharging is not profitable.

It can be shown that when both accept cards and surcharge, each merchant earns profits of $\pi = \frac{1}{2}t - K$. This is the same as when both merchants accept cards and do no surcharge, and a proof proceeds along the same lines as in the no-surcharging case. The next lemma identifies the values of b_S for which deviating by not accepting cards is not profitable.

Lemma 4 (Merchant acceptance threshold with surcharging)**.** *If both merchants accept cards and surcharge, neither merchant is better off not accepting cards provided that*

$$b_S \geq m - E[b_B|b_B \geq m] + \frac{3t(1 - \sqrt{1 - \frac{2K}{t}}}{1 - G(m)} \tag{16}$$

The acceptance threshold is nearly identical to that in the no-surcharging case, even though the problem facing merchants is different in the two cases. There are two differences. First, $E[b_B|b_B \geq m]$ appears instead of $E[b_B|b_B \geq 0]$. Since this term is increasing in m, this tends to lower the benefit threshold above which merchants accept cards. Second, as m increases, $1-H(m)$ decreases, and average fixed costs increase, thus reducing merchants' inclination to accept cards. It is somewhat surprising, though, that the first term survives unaltered. Intuitively, it might seem merchants should be indifferent to the level of the merchant fee, since they can choose to pass it on to consumers. The model shows this intuition to be erroneous.

Establishing the conditions under which a merchant will not want to deviate by accepting cards and not surcharging is more involved, and it does not result in a simple closed-form expression, which can easily be interpreted. A derivation of the no-deviation condition in this case can be found in the appendix. I examine merchant's choice problem more closely as part of the numerical analysis of the model. That analysis shows that there might be multiple, even asymmetric, equilibria depending on the values of m and b_S. For some values there will actually no be any equilibria. This makes it impossible to derive optimal merchant fees is in the same fashion as in the no surcharging case.

A simpler problem is to examine what were to happen if the only were alternatives were not accepting card and accepting cards and surcharging. In the following I abstract from merchant fixed costs in order to simplify the analysis. If merchant fixed costs were included, it would entail a negative effect on card usage, at least assuming that merchant fees are positive. That would be the case because the fixed cost element in merchants' decision problem increases as the fraction of card using consumers shifts from $1-H(0)$ to $1-H(m)$, contributing to making merchants less willing to accept cards. Thus, ignoring fixed costs understates any negative effects surcharging might have on card usage.

With surcharging the banks and the social planner must take into account the fact that higher merchant fees reduce consumers' willingness to use cards, a feature of the problem which was absent in the no-surcharging case. If one solves the banks' optimization problem, the optimal merchant fee is indirectly defined by

$$m = c_C - c_D - \frac{\mu(m)}{\mu'(m)} \tag{17}$$

At first glance, this looks like the corresponding condition when surcharging is prohibited. However, where $\mu'(m)$ equals $-(1 - H(0))g(m - E[b_B|b_B \geq 0])$ in the no-surcharging case, the expression with surcharging becomes

$$\mu'(m-\delta_m) = -h(m)(1-G(m-E[b_B|b_B \geq m])) - (1-H(m))(1+mh(m))g(m-E[b_B|b_B \geq m]) \tag{18}$$

Solving the social planner's problem likewise gives a more complicated expression for the optimal merchant fee, and it is no longer a simple function of costs. An effect of surcharging, which can be established, is that on card usage.

Proposition 4. *If merchant fees are positive and merchants surcharge fees when permitted to, card usage is lower under surcharging than no surcharging.*

As a corollary, note that the opposite occurs when merchant fees are negative. In that case card usage increases if merchants surcharge (i.e. pass the rebate on to consumers).

4.2 Banks face fixed costs

In the preceding, I assumed that the costs of processing and handling cash and card payments were proportional to the usage of cards and cash in the economy. A more realistic assumption might be that there are some fixed costs involved in operating a card payment system. This would be consistent with the evidence of decreasing average variable costs of card payments, which are document in (Schmiedel et al., 2012).

Assume now that banks must pay a fixed cost F associated with producing card payments. This alters their rationality constraint which now reads

$$\mu(m)(m - c_C) - c_D(1 - mu(m)) - F \geq -c_D \tag{19}$$

It is straightforward to verify that this constraint is no longer satisfied at the socially optimal fee. Inserting the unconstrained optimal fee into banks' objective function, it evaluates to $\Pi = -c_D - F$, which is less than the banks' profits in a pure cash economy, which are $-c_D$. Instead, the social planner must impose a constrained optimal fee, which is given in the following lemma.

Lemma 5 (Constrained socially optimal merchant fee)**.** *The socially optimal merchant fee equals to difference in marginal cost between card and cash payments plus the average fixed cost of card payments given that merchant fee.*

$$m = c_C - c_D + \frac{F}{\mu(m)} \tag{20}$$

The addition is that banks must be compensated to cover their fixed their average fixed costs. Note that the equation only defines the merchant fee indirectly as it appears on both sides of question. To establish the conditions under which there exists an m that satisfies the equation, one can appeal to a number of fixed point theorems in order to identify the restrictions which must be satisfied.

The banks' profit-maximizing fee does not depend on the presence of a fixed cost; I will assume that banks' rationality constraint is satisfied at this fee, since otherwise there would exist no fee at which banks would be willing to produce card payments. Just glancing at the expression for the privately and socially optimal fees and comparing them, it is not immediately obvious that the privately set fee must exceed that of the social planner. Some consideration shows that this must be the case. At the privately optimal fee, banks' profits exceed those earned in a pure cash economy by assumption. The unconstrained socially optimal fee is less than the privately optimal fee, and the social objective function is decreasing for m greater than the unconstrained social optimum. That implies that fees can be reduced below the privately optimal fee and still satisfy banks' rationality constraint, whereas any fee higher than the privately optimal fee would be associated with lower welfare.

From a regulator's perspective, the expression for the constrained socially optimal fee is less satisfactory. Identifying the optimal fee imposes a greater informational demand as the optimal fee depends on knowledge of the distribution of merchant benefits (as well as the parameters t and K if merchant fixed costs are taken into account). The analysis, however, points to a superior solution from a welfare perspective, namely to impose the unconstrained optimal fee and reimburse the banks the fixed costs of operating the card payment system, since that does not involve a welfare loss from reduced card usage.[9]

4.3 Incentives to innovate

A policy issue, which is somewhat related to the fees that banks charge or are permitted to charge, is innovation. Innovation is inherently difficult in payment schemes as multiple parties are involved in the production of a payment and must therefore adapt simultaneously. For instance, an improvement in card security might require consumers to get obtain new cards and merchants to invest in improved card terminals. The problem is exacerbated in the case of entry to the market, where potential entrants confront a hen-and-the-egg problem in persuading consumers and merchants to use the scheme, and in addition may face resistance from banks who control the payment infrastructure. Setting aside such difficulties, there is the problem of financing innovation and whether banks face the correct incentives to invest. In this section I briefly discuss that issue.

In the context of the model, a distinction can be made between innovations that reduce costs, and those which increase cardholder benefits or merchant benefits. It would seem that banks should agree with a social planner on the value of reducing costs, as banks fully bear this cost. It is less obvious that this is the case for the benefits enjoyed by consumers and merchants. This problem is likely to be particularly pronounced when the socially optimal fee is applied.

This claim about costs can be verified by comparing the objective functions of banks' and the social planner. To gauge the value of a cost reduction, one can differentiate the respective value functions of the banks and the social planner. Differentiation with respect to c_C affects the profit function (or social planner's objective function) directly, an effect of magnitude $-\mu(m)$, and indirectly, since the optimal merchant fee is itself a function of costs. However, the latter effect is zero by the envelope theorem.

To illustrate the point that banks might invest insufficiently in innovations that enhance benefits, consider a investment that would increase $b_{S,max}$. The marginal value of this increase, again applying the envelope theorem, is

$$\frac{\partial \Pi}{\partial b_{S,max}} = g(b_{S,max})(1 - H(0))(m_{opt} - c_C + c_D) \tag{21}$$

If the unconstrained socially optimal fee were imposed, this would evaluate to zero. If the social planner were to ensure that the banks' individual rationality constraint remain satisfied, the social planner should allow the banks to additionally charge the average fixed cost of the cost of the innovation, in exactly the sammer manner as the constrained socially optimal merchant fee includes an average fixed cost component.

The above expression can be contrasted with the marginal value that the social planner would assign to an increase in $b_{S,max}$

$$\frac{\partial W}{\partial b_{S,max}} = g(b_{S,max})(1 - H(0)\,(E[b_B|b_B \geq 0] + b_{S,max} - c_C + c_D) \tag{22}$$

[9] The Danish national debit card, the Dankort, is actually financed in a manner, which bears similarity to this. Merchants do not pay transaction fees, but rather an annual subscription that is calculated ex post, i.e. once the total number of transactions is known, and is set to cover half the costs of producing the payments.

To give an example of the difference in incentives facing banks and the social planner, the case of uniformly distributed benefits can again be considered. We found earlier that $m_opt = \frac{1}{2}\left(c_C - c_D + b_{S,max} + \frac{1}{2}b_{B,max}\right)$ and $(E[b_B|b_B \geq 0] = \frac{1}{2}b_{B,max}$. Inserting these into the above expressions and rearranging, one sees that

$$\frac{\partial W}{\partial b_{S,max}} = 2\frac{\partial \Pi}{\partial b_{S,max}} \tag{23}$$

That is, the social planner values an increase in merchant benefits twice as much as do banks. More generally, comparing the objective functions of banks' and the social planner points the sources of bias in banks' incentives. In the case of increases in merchant benefits, there are two sources of bias. As $b_{S,max}$ increases, more merchants accept cards. This benefits all of the consumers, which prefer to use cards, a positive externality which banks do not take into account in their decision making. The second source of bias is that banks' marginal revenue from the increase in card transactions is proportional to m_opt, whereas the corresponding value in the merchant's expression is $b_{S,max}$.

Similar considerations hold for increases in consumer benefits. The calculations are essentially the same in the case of the banks, only with the roles of the distribution functions reversed. For the social planner the problem is slightly more complicated as a third effect must also be taken into account. As $b_{B,max}$ increases, so does the average benefit to card holders $E[b_B|b_B \geq 0]$. This means that more merchants will accept cards as well. This is likely to reduce welfare, since the marginal merchant might enjoy negative benefits.

4.4 Pricing of payment instruments

The key assumption of the model is that banks rarely charge consumers card transaction fees. Empirical evidence on this matter can be found in e.g. the national cost of payment studies, on which the European Central Bank (Schmiedel et al., 2012) study is based. A number of these studies have been published, and some show that banks not only do not charge transaction fees, but frequently fail to even cover their costs, particularly in the case of cash payment services. This type of price-setting behavior is consistent with the observation that banks set prices differently depending on the type of banking product, with some products being bargains and others rip-offs (Heffernan, 2002).

A possible explanation of this is that banks use free payment services as a means to attract retail consumers who provide inexpensive funding in the form of the deposits.[10] The Economist describes banks' pricing of payment instrument as follows: "To consumers, most payments appear to be free because they are given away by banks as part of a bundle of banking services that some customers subsidise through low interest rates on deposits."[11]

While the primary aim of this paper is not to rationalize banks' underpricing of payment services, but rather to examine its consequences, a heuristic argument may provide some motivation. Consider a simple one-period model involving a consumer and a bank. The consumer has savings, which can be deposited with the bank at time zero. The consumer has an outside option, which can be thought of as either using cash and not using the banking system or perhaps going to a different bank. The savings are to be spent in period one where the consumers pays for consumption either using a payment card or by a cash withdrawal and subsequent cash payment. The bank has costs c_{card} and c_{cash} associated with producing the card payment and the cash withdrawal. Assume, as in the related literature such as Rochet and Tirole (2002), that consumer banks are imperfectly competitive, but that this manifests itself in the spread between deposit and lending rates rather than in a markup on payment services. Consumers are therefore

[10]The Danish cost study (Danmarks Nationalbank, 2011) reveals that banks fail to recover their costs for practically all of the payment services they provide. However, the authors of the study also point out that banks' deficit all but vanishes when taking into account the low interest rates paid on current account balances.

[11]The Economist, 19 May 2012.

valuable to banks because they provide cheap funding. Denote this value δ, and consider the case of $\delta > c_{cash} > c_{card}$. Finally, assume that the consumer enjoys card benefits b_B, which are unobservable to the bank. These benefits are distributed with cumulative distribution function $H(b_B)$. The bank's problem is to set prices p_{cash} and p_{card} for the use of payment services.

It is then possible to specify the conditions under which an equilibrium with $(p_{cash}, p_{card}) = (0, 0)$ results. To construct such an equilibrium, consider first a consumer for whom $b_B < 0$. Such a consumer will deposit his savings with the bank if $p_{cash} \leq 0$; otherwise the outside option of not using the bank is more attractive. There is also the possibility that $p_{cash} < b_B < 0$. Assume for now that this is not the case; in a moment the conditions under which this is the case will become clear. Assume for the moment that the latter is not the case. It then makes sense for the bank to set cash withdrawal prices at exactly zero. Now, assume the price of card payments is also zero and ask whether it will make sense for the bank to deviate from this price. A positive price will entail extra revenue from remaining consumers of $(1 - H(p_{card}))p_{card}$ and a loss of $(H(p_{card}) - H(0))(\delta - c_{card})$ from departing consumers. If δ is large relative to c_{card} or if consumers are price sensitive, a positive price will result in a loss.[12] What about a negative price? This is a possibility since it may result in savings from cash users of $(H(0) - H(p_{card}))(c_{cash} - c_{card})$. On the other hand, it will involve an additional cost of $(1 - H(0))p_C$ because the rebate would also benefit consumers who would otherwise use cash. Thus, if the costs of providing cash and card payment services are not too different, free provision of payment services is an equilibrium outcome that is consistent with, even a result of, imperfect competition.

The above is intended only as a suggestive of why underpriced payment services might be an equilibrium outcome, and it is does not explain the preponderance of flat-fee contracts[13]. A more realistic model might feature elements such as a model of bank competition, switching costs and asymmetric information about not just customer preferences, but also their potential value to banks. It does, however, capture some salient features of typical bank contracts. For instance, bank contracts typically specify that payment services are only free provided that customers provided that certain conditions are met, for instance that consumers maintain their salary accounts with the bank, or that customers have a certain volume of business with the bank. An exception is typically made for younger customers, likely reflecting asymmetric information about the value of such consumers to the banks.

5 Policy Analysis in a Calibrated Model

In this section I first discuss data on payments and how these can be used to calibrate the parameters of the model. As an application, I compare the optimal interchange fee produced by the model with interchange fees determined by alternative methods. Finally, I analyse the model when surcharging is permitted. That final part of the analysis is intended only to be illustrative of the the qualitative effects of surcharging, since to analyze the model with surcharging assumptions on the distribution of consumer and merchant benefits are required,. I examine the case of uniformly distributed benefits.

The data used for calibrating the model parameters is from a study of the cost of producing payments conducted by the Danish central bank (Danmarks Nationalbank 2011). This study was conducted as part of a European Central Bank project involving 13 national central banks. A summary of the results is given by (Schmiedel et al., 2012). The Danish data are particularly well-suited for calibrating the model for at least two reasons. First, the published data are detailed

[12] I assume that consumers with a preference for cards use the outside option rather than remain with the bank, using cash withdrawals.

[13] DellaVigna and Malmendier (2006) discuss various reasons why consumers might prefer flat-fee contracts in general. Since their application is gym membership decisions, some of the possible arguments are unlikely to carry weight in the context of financial services.

enough to perform a mapping of the data to the model parameters; the (Schmiedel et al., 2012) data, in contrast, is published at a higher level of aggregation. Second, the assumptions of the models are satisfied in the market. There is a single card, which enjoys what is essentially a monopoly position, and consumers with few exceptions do not pay transaction fees.

From the cost study I draw on 2009 data relating to the use and cost of producing of cash and Dankort payments. The cost data have further been split into fixed and variable costs in a related analysis by Jacobsen and Pedersen (2012). The Dankort[14] is a national debit card, and more than 90% of card payments in Denmark involve the Dankort. The fee structure of the Dankort is unusual by international standards. Merchants do not pay regular merchant fees, nor is there an interchange fee. Instead, merchants pay an annual subscription fee, which is set such that merchants' collective subscription payment covers half of banks' production costs. According to survey data, roughly 95% of all merchants accept Dankort payments. Consumers generally do not pay for the use of the Dankort. Even though list prices often include an annual fee, these fees are typically waived if e.g. consumers have their current account with the bank or a certain amount of other business with the bank. Indeed, the cost study shows that banks' total Dankort income from consumers is negligible.

The intention of the analysis is not to argue that the results will generalize to other markets, but to illustrate an application of the model. The results evidently depend on the specifics of a market. As an example, Danish labor costs are high whereas technology costs are low. The result is a low, even negative optimal interchange fee. A similarly low fee is an unlikely outcome in a number of other countries where, for example, cash handling costs are lower due to lower labor costs. Still, unless one is willing to argue that card payments are substantially more expensive to produce than cash payments - and those who favor higher interchange fees tend to make the opposite argument - the model will produce low interchange fees across markets.

5.1 Calibration of model parameters

The key parameters of the model for the purpose of calculating optimal interchange fees are c_C and c_D. These reflect banks' marginal costs as the fraction in the economy of card and cash transactions change. In the model, these figures refer to social costs, and for simplicity I assumed that banks' private costs were equal to these when, in fact, the private costs may be different as banks do obtain some income, e.g. in the form of periodical fees, from consumers. This income is rather limited, however. Other model parameters that appear in derivations or results include merchant's production cost (γ), distance costs (t), the fixed costs involved in accepting (K) and (k) adopting cards, and finally the distributions of consumer (b_B) and merchant benefits (b_S). Some of these parameters, notably the cost parameters, can be estimated based on the cost study data. Others such as the distribution of consumer benefits require a greater degree of judgment. The key model parameters, not including k which is discussed later, are summarized in table 1.

[Table 1 about here.]

The model parameters are obtained as follows. Thinking of the model as describing a single transaction, the choice of γ is essentially a choice of transaction size while the choice of t determines the profit margin. According to the Danish cost study (Danmarks Nationalbank, 2011), the average transaction size is 257 Danish kroner[15]. Since the merchant price is close to $\gamma + t$, I choose estimates of $\gamma = 230$ and $t = 27$, corresponding to a markup of slightly above 10%. In the model, K corresponds to the total fixed cost, but the model describes a situation in which there is a unit mass of transactions. The logical counterpart when thinking of a single transaction is to compare the merchant fixed costs to the total value of cash and card payments. The total

[14] I also include the co-branded Visa/Dankort in this category.

[15] All units, unless otherwise stated, are in Danish kroner. The currency is pegged to the euro at an exchange rate of about 1 euro = 7.43 kroner

value of cash (150.5 billion) and Dankort payments (253.1 billion) is kr. 403.6 billion according to table 1 in the cost study, while merchant fixed costs are 756.9 million according to (Jacobsen and Pedersen, 2012). Part of this cost, however, is fees of 245.7 million paid to banks, which must be subtracted. For an average payment of size 257, this produces a value of $K = 0.326$.

The parameters c_C and c_D reflect banks marginal costs of cash and card payments. Since banks' production functions are not observable, the most obvious proxies for these figures are the average variable costs. The use of this proxy is correct in the case of constant marginal costs. The empirical evidence presented in Schmiedel et al. (2012) suggest that this may not be an unrealistic assumption.[16] According to Jacobsen and Pedersen (2012), issuing banks' variable cash-related costs are 894.2 million. The total variable costs for issuing and acquiring banks are somewhat harder to obtain as one cannot simply add their individual costs. This is because of fee payments between issuers and acquirers, which would entail some double-counting. If one corrects for such fees, the total variable cost of issuers and acquirers is 316.3 million.[17] In order to identify banks' private costs, I correct for fee income to banks. Banks' fee income from providing cash services is 373 million, and their fee income from households for the Dankort is 81.6 million. Dividing the total average variable cost net of fee income by the total value of respectively Dankort and cash transactions and then multiplying by 257, I arrive at average variable net costs of 1.527 (social cost) and 0.890 (private cost) in the case of cash and 0.321 (social cost) and 0.238 (private cost) in the case of cards. As an aside, the low fee income of from consumers is consistent with the claim that banks' derive limited revenue from payment services. There were 4.2 million issued Dankort in 2009, so the average annual fee income per card was less than 20 Danish kroner (less than 3 euros).

The more difficult parameterizations are those involve the distributions of benefits. Merchant benefits b_S can be interpreted as the marginal savings (or cost) from a card payment relative to a cash payment. A reasonable estimate of b_S is 0.5825. I arrive at this figure by calculating the difference in merchants' marginal cost of receiving cash and cards for a payment of average size.[18] I treat the estimate as the midpoint of a uniform distribution and thus let $b_{S,min} = 0.7 - x$ and $b_{S,max} = 0.7 + x$. A distribution for b_B can be determined in the same way. According to the cost study, consumers have net costs of 0.01 for cash and 0.003 for Dankort per krone spent. For a transaction of average size this means that consumers enjoy benefits of 1.8 by paying with Dankort. I therefore set $b_{B,min} = 1.8 - y$ and $b_{B,max} = 1.8 + y$.

The values of x and y are chosen such that the fraction of merchants accepting Dankort and the fraction of consumers preferring to pay by Dankort match real figures. About 95% of all merchants accept payments by Dankort. Since the total fraction of card payments (by value) is 62.7%, the fraction of consumers preferring to pay using Dankort is set at 66% (= 0.627/0.95). y can then be found first by solving the following equation for y

$$1 - H(0) = \frac{b_{B,max}}{b_{B,max} - b_{B,min}} = \frac{b_{B,mid} + y}{2y} = 0.66 \tag{24}$$

Since $m = 0$ in the case of the Dankort, the merchant threshold is

$$b_S \geq b_S^{\star} = -E[b_B | b_B \geq 0] + \frac{3t(1 - \sqrt{1 - \frac{2K}{t}})}{1 - H(0)} \tag{25}$$

[16] See, for instance, chart 2 in the paper, which depicts the relationship between average unit costs and payments per capita. The lines are downward sloping at a decreasing rate, which is consistent with how unit costs would evolve if marginal costs were constant.

[17] This figure cannot be directly gleaned from the cost study or related papers. I have obtained the underlying data from the authors to calculate the figure

[18] Figures for merchant's variable costs are provided in (Jacobsen and Pedersen, 2012), which itself is based largely on a working paper (only published in Danish) by Jacobsen. According that paper the marginal cost functions for cash and Dankort payments can be written as $mc_D = 0.70 + 0.0027V$ and $mc_C = 0.76 + 0.0002V$, where V is the transaction value. The estimate is found by inserting $V = 257$, and taking the difference between the marginal cost estimates for cash and cards.

This can be calculated based on the found parameter estimates, and x is then set so as to satisfy

$$1 - H(b_S^\star) = \frac{b_{S,max} - b_S^\star}{b_{B,max} - b_{B,min}} = \frac{b_{S,mid} + x - b_S^\star}{2x} = 0.95 \tag{26}$$

The obtained parameter estimates indicate that there is more heterogeneity in consumer benefits than merchant benefits. Another point of interest is the merchant acceptance threshold implied by the model, which evaluates to -2.22, implying that a merchant would accept card payments even if it increased costs relative to receiving payment pay cash by 2.22 for an average payment or by 0.86% in percentage terms.

5.2 Optimal interchange fees

The unconstrained optimal merchant fee follows directly from the parameter estimates. It is

$$m_{opt} = c_C - c_D = 0.321 - 1.527 = -1.206 \tag{27}$$

Since the model has been calibrated based on an average transaction of size 257, in proportional terms this corresponds to an optimal merchant fee of -0.47%. Total acquiring costs for the Dankort, corrected for (i.e. not including) fees by paid acquirers to issuers, are estimated at 87 million[19], which corresponds to acquiring costs of just 0.03%. Hence, the implied optimal interchange fee is -0.50%. What informs such a low, even negative fee? Recall that $m_{opt} = c_C - c_D$. The main explanation for the low fee is to be found in banks' high net costs of providing cash services, c_D. Banks' revenues from cash services are negligible, but their costs are substantial. This effectively means that banks are subsidising the use of cash, and it is this tendency that a social planner seeks to counteract by imposing a negative interchange fee. Some further analysis, however, shows that the optimal fee is not the only optimal fee; in derivations, I assumed an interior solution, but in practice all merchants accept cards at the optimal fee, and at higher fees as well (see figure 2). Imposing such a low fee is therefore not the only way of reaching the optimal fee: A whole interval of fees would accomplish this task.

It is also possible to calculate the privately optimal interchange fee as predicted by the model given the calibrated parameter estimates. In the case of no fixed costs, for instance, the privately optimal merchant fee is 1.21%. While this figure relies on the arbitrary assumption of uniformly distributed benefits, it is not too unrealistic when compared to actual fee levels.

This fee can be contrasted with fees obtained by alternative methods. It is, in general, not transparent how banks or card networks set interchange fees. However, a common method appears to be that issuers must be compensated for bearing certain costs that benefit the merchant side of the market (Börestam and Schmiedel, 2011). These costs include, as a minimum, operating costs related to processing payments, guaranteeing payments, and other security costs. An alternative, which some regulators use as a benchmark, is to apply the merchant indifference test. Based on the cost study data estimates of interchange fees based on both of these methods can be obtained.

The cost-based interchange fee depends on which issuer costs are assigned to the merchant side of the market. The total cost of issuer services for the Dankort is estimated to be 591.3 million[20]. Of these, 117.2 million are purely related to transactions processing and guaranteeing payments. If these costs were to be covered, an interchange fee of 0.05% would be required. If all issuer costs were to be covered via the interchange fee, the implied fee would equal 0.23%, still a very low figure compared to actual interchange fees. Note that these calculation assume the total value of Dankort payments to be fixed. The model predicts that fewer payments would take place so presumably a higher interchange fee would be required.

[19]This estimate is based on data collected as part of the Danish cost study, but which cannot be directly found in the publication. I have obtained the more granular data from the authors.

[20]Some of these costs are actually borne by the acquirer (there is only one acquirer of the Dankort), which in the Danish market also performs some services that are traditionally performed by issuers

The low interchange fee implied by the cost-based method is due mainly to two factors. First, the number of Dankort payments is large. Card usage per capita is high in Denmark, and the vast majority of card payments are made using the Dankort. Second, the cost of producing Dankort payments is low. For instance, issuing banks' cost of producing Dankort payments is about 2.3 times their cost of producing payments with international debit cards even though the number of Dankort payments is roughly 20 times greater. If a similar cost-based analysis were to be performed for international debit cards, the implied interchange fee would be in the range 0.4%-0.7% depending on which issuer costs are passed on to the merchant side of the market.

A final alternative is the merchant indifference test, which calls for an interchange fee that makes the merchant no worse off by receiving payment by card. A practical challenge confronting this test is that merchant's marginal cost depends on the transaction value. However, the following marginal cost functions for cash and Dankort can be derived from the cost data in Jacobsen and Pedersen (2012)

$$mc_D = 0.70 + 0.0027V \tag{28}$$

$$mc_C = 0.76 + 0.0002V \tag{29}$$

It follows that that way to make a merchant indifferent is to impose an interchange fee a_{id} of

$$a_{id} = -0.06 + 0.0025V \tag{30}$$

This implies that an interchange fee of 0.25% less a fixed payment of 0.06. Again, similar calculations could be performed for international debit cards for which merchants' marginal cost function is estimated to be

$$a_{id} = 0.75 + 0.0051V \tag{31}$$

The variable part of the marginal cost includes the existing interchange fee, the level of which is not known. However, an upper bound estimate can be obtained by assuming that all of merchants' fees paid on international debit cards (31.5 million according to table 5.7 in Danmarks Nationalbank (2011)) are interchange fee payments. Since the value of international debit card payments was 6.8 billion, an estimate of the maximum average interchange fee is 0.46%. This would imply a marginal cost function, exclusive of interchange fees, of $a_{id} = 0.75 + 0.0006V$. It follows that the optimal interchange fee according to the merchant indifference test should therefore be of the order

$$a_{id} = -0.05 + 0.0021V \tag{32}$$

Interestingly, this implied interchange fee level is close to the interchange fee of 0.20%, which the European Commission has agreed upon with Visa and MasterCard for cross-border debit card transactions in Europe.

When calculating the value of the social planner's objective function, it turns out that the choice between different benchmarks for interchange fees may not make much of a difference. The reason is that for low enough merchant fees, the merchant acceptance threshold may lie below the numerical estimate for $b_{S,min}$. In that case all merchants accept cards, and reducing the merchant fee further has no welfare consequences. When fixed costs are included, the value of m at which this occurs is -0.32 (-0.12%), but if fixed costs are excluded the value is 1.17 (0.46%). Below these values the social planner's objective function takes on a constant value. Merchant profits and the value of the social objective function as a function of the merchant fee are depicted in figure 2.

[Figure 2 about here.]

5.3 Surcharging

When surcharging is permitted, the model does not provide easy answers to whether merchants will, in fact, surcharge. The numerical analysis highlights this difficulty. To illustrate this with an example, the equilibrium outcome for a pair of average merchants ($b_S = 0.5825$) is to accept cards and not surcharge if $m = -1$, to accept cards and surcharge if $m = 2$, not to accept cards if $m = 5$. This might indicate a natural progression from not surcharging at low fees, to surcharging at higher fees, and then to not accept cards when fees reach a certain limit. Such a conclusion does not hold for all combinations of m and b_S, however. When $(m, b_S) = (-0.2, -2.3)$, for instance all three symmetric outcomes are equilibria. If $(m, b_S) = (4.7, 2.6)$, there are two asymmetric equilibria in which one merchant accepts cards and surcharges and the other does not accept cards. Increase b_S slightly, say to 2.66, and there are no equilibria, whereas both merchants surcharge if $b_S = 2.7$.

At the socially optimal fee derived under the assumption of surcharging being prohibited, the equilibrium is for all merchants with $b_S \leq 0.70$ is to accept cards and not surcharge. However, for merchants who enjoy greater benefits the equilibrium outcome is an asymmetric one in which one merchant surcharges and the other does not. The implication is that at that fee, all merchants accept cards, and some merchants surcharge whereas others do not. At somewhat higher fees such as 0.20% (0.54 in absolute value), which is close to the fee according to the merchant indifference test, most merchants ($b_S > -1.83$) accept cards and surcharge. Merchants who enjoy lower benefits do not accept cards at all. At the privately optimal fee (i.e. that fee that is optimal when surcharging is not permitted), merchants do not accept cards for $b_S \leq 0.63$, but accept cards and surcharge when $b_S > 0.72$. At a small interval in between there are two equilibria in which one merchant does not accept cards, and the other accepts cards and surcharges.

An overview of the equilibria of the model is provided in figure 3. It depicts the regions of m, b_S-combinations for which different types of equilibria in the merchant game occur. The dark regions indicate the presence of equilibria.

[Figure 3 about here.]

Merchants do not accept cards if merchant fees are high and they enjoy low benefits. When merchant fees are positive, merchants will generally surcharge if permitted, though there are also cases with no or asymmetric equilibria. When merchants are slightly negative, merchant who enjoy the lowest benefits will still surcharge. The predominant equilibria when merchant fees are negative, though, is to accept cards and not surcharge in the case of low merchant benefits or for one merchant to surcharge and for the other not to surcharge.

To gauge welfare in the surcharging case against welfare in the no-surcharging case, one can evaluate the value of the social planner's objective function under the assumption that merchants face only a choice between not accepting cards and accepting cards and surcharging. The presence of e.g. situations without any equilibria makes it impossible to evaluate the objective function when all of merchants' three options are available.

[Figure 4 about here.]

If one compares welfare under surcharging and no surcharging, the welfare effects of surcharging depend on the merchant fee applied. For negative merchant fees welfare is greater when surcharging is applied. This is because surcharging of negative fees, which is equivalent to giving rebates for card usage, induces more consumers to use cards, which benefits card-accepting merchants who, on average, benefit from greater card usage. It may not be obvious from visual inspection of figures 1 and 3, but the opposite is the case for positive merchant fees for which welfare is higher when surcharging does not take place.

Comparing figure 4 to figure 2 also shows that banks' privately optimal fee is lower when surcharging takes place than when it does not, as are bank profits. An interesting issue is how social compares at the privately optimal fees with and without surcharging. On the one hand, we know that surcharging lowers welfare for a given (positive) merchant fee, but on the other hand the privately optimal merchant fee is lower when surcharging is permitted. It is at least plausible therefore that surcharging will increase welfare. In fact, this turns out to be the case, though the improvement in welfare is small. In the model with merchant fixed costs, the value of the social planner's objective function is 1.22 when surcharging is applied and 1.17 when merchants do not surcharge. The comparable figures in the model without merchant fixed costs are 1.70 and 1.69.

The implication is that surcharging might improve welfare if merchant or interchange fees are not regulated. At low interchange fees, however, welfare would be greater in the absence of surcharging. While welfare would in principle be higher with surcharging for negative fees, it is not clear that merchants would surcharge in that case as illustrated by figure 3.

5.4 Consumer adoption costs

When both consumers and merchants face adoption costs there will, in general, be two card usage equilibria, one with high adoption among merchants and consumers and one with low adoption. The first step in identifying these equilibria is to select a value for consumer adoption costs k. I use an estimate is $k = 0.5$.[21] Given that value it is possible to find equilibria for different values of m. For reasonable values of m there will be two equilibria. The numerical analysis shows that both bank profits and the value of the social planner's objective function are always highest in the equilibrium where card usage is high. The merchant fee that maximizes profits in that equilibrium is $m = 2.68$, or 1.04% as a percentage of the transaction value. The fixed points of the function $f(b_S)$ are depicted in figure 5 for three different values of the merchant fee.

[Figure 5 about here.]

This figure is typical of how the two fixed points are located. To understand the figure, consider first the case of b_S being lower than at the first equilibrium point. A low value of the merchant threshold b_S corresponds to many merchants accepting cards. In that case, many users would also like to use cards. The marginal card user therefore attaches a low value to using cards relative to cash, but in that case merchants are less inclined to compete for card users. So for the value of the consumer threshold, i.e. the value of b_B that is consistent with the particular value of b_S, $f(b_S) > b_S$, which means that fewer merchants will accept cards. As b_S increases in value, two effects are at play. First, the marginal card user attaches more value to being able to pay by card, and merchants compete more vigorously for these with a resulting decrease in $f(b_S)$. This effect dominates initially. As there are fewer card users, however, merchants' fixed cost must be spread across fewer users, and ultimately this effect takes over.

In the high card usage (low b_S) equilibrium, 58.0% of consumers use cards and 55.9% of merchants accept cards. If merchants and consumers were to find themselves in the low card usage equilibrium, the fraction of merchants and consumers accepting and using cards would just be 8.5% and 13.8%. These figures contrast with merchant acceptance at the fee, which maximizes the social planner's welfare function assuming that consumers and merchants end up in the high card usage equilibrium. That fee is -0.17, and is associated with 61.6% of consumers using cards and all merchants accepting cards.

The model is static, but if one were to think of the model in dynamic terms, only the equilibrium with high card usage would be a stable equilibrium. If, initially, the merchant acceptance threshold were less than at the first equilibrium point, the number of merchants accepting cards would decrease due to the mechanics described above: Many consumers would use cards, and

[21] According to survey data, a typical consumer makes about 400 transactions per year. Supposing that banks did charge consumers their official list prices of 200 kroner, that would correspond to a value of $k = 0.5$ if such prices were charged, since k can be viewed as a consumer's average fixed cost.

this would actually make it less attractive for merchants to accept cards, corresponding to an increase in the merchant threshold. The opposite would happen for values of b_S between the two equilibrium points. Few consumers would be using cards, but since these consumers would be those enjoying high card benefits, more merchants would accept cards in order to attract them and the merchant acceptance threshold would decline. To the right of the second equilibrium point the situation would diverge towards an outcome without any card usage.

6 Conclusion

This paper a number of contributions to the literature on interchange fees. As general matter it shows that many conclusions of the earlier literature are altered, sometimes entirely reversed, when one entertains the assumption that consumers' marginal cost of paying by card is zero. One implication is that lower interchange lead to greater card usage. The intuition is that the absence of transaction fees makes the market essentially one-sided. Higher interchange fees reduce card usage as fewer merchants accept cards, but there is no counterbalancing effect on the consumer side of the market. A related conclusion is that banks effectively act as monopolists when setting the interchange fee. The outcome is interchange fees which exceed the social optimum and a price structure which results in an excessive use of cash. These observations may help explain the observation that card usage is particularly high in countries in which card schemes have operated for years without interchange fees.

In the baseline model the optimal interchange fee depends solely on the relative costs of producing cash and card payments. This feature distinguishes it from other models in which optimal interchange fees depend on the distribution of benefits that consumers and merchants derive from card payments. Such a results useful from a regulatory point of view since costs can be calculated whereas consumer benefits are unobservable. While the existence of a purely cost-based interchange fee fails if the model's assumptions are altered, the cost-based optimal fee still provides a useful benchmark for calculating optimal interchange fees. As an example, if one includes merchant fixed costs in the model, the socially optimal fee can be shown to equal the cost-based fee and another term, which is negative. In that case the cost-based fee therefore provides an upper bound for the socially optimal fee.

Due to the presence of a cost-based fee, I am able to compute an optimal interchange fee based on data from Danmarks Nationalbank (2011), which is part of the ECB cost of payments study. The resulting optimal interchange fee is negative. It is so because banks' pricing behavior, which effectively subsidizes cash usage. The optimal response for a social planner is require a low cost of card payments. I also compare the obtained interchange fee level to the fee level derived from both a method akin to that used by industry and the level based on merchant indifference or tourist test, which some regulators use as a benchmark. The fee based on the merchant indifference test is slightly higher than the optimal fee implied by the model, but still low compared to the privately optimal fee.

The model shows the effects of surcharging to be ambiguous. At low, but positive merchant fees welfare is greatest when merchants do not surcharge. While welfare would be highest if merchant fees were negative and merchants did surcharge, it is not clear that merchants would surcharge given such fees. The model predicts that merchants would mainly choose to surcharge when fees are positive. If merchant or interchange fees are unregulated, surcharging might improve welfare. While surcharging reduces welfare for positive merchant fees, it also reduces privately optimal fees.

Appendix

Proof of lemma 1

The proof of lemma 1 mirrors the proof in Rochet and Tirole (2002), except for the fact that fixed adoption costs are included in this derivation. The inclusion of the proof is useful for an additional reason, namely that the derivation of surcharging equilibria follows a similar structure, but a number of intermediate calculation steps are not shown in that derivation.

Assume initially that both merchants accept cards, and set prices p_1 and p_2 in order to maximize profits. For a merchant pair, let x denote the market share of one of the merchants, say merchant 1. In that case, the market share is found by solving

$$p_1 + xt = p_2 + (1-x)t \Leftrightarrow x = \frac{1}{2} + \frac{p_2 - p_1}{2t} \tag{33}$$

The resulting merchant profits are

$$\pi_1 = \left(\frac{1}{2} + \frac{p_2 - p_1}{2t}\right)(p_1 - \gamma - (m - b_S)(1 - H(b_B^\star)) - K \tag{34}$$

$$\pi_2 = \left(\frac{1}{2} + \frac{p_1 - p_2}{2t}\right)(p_2 - \gamma - (m - b_S)(1 - H(b_B^\star)) - K \tag{35}$$

In words, profits are the product of market share times profit margin, less fixed costs. Note that merchant fees and card benefits are only paid and received for the fraction $1 - H(b_B^\star$ of consumers, which prefer to pay by card. Recall that $b_B^\star$ denotes the threshold above which consumers prefer the use of payment cards to cash. This is assumed to be zero throughout most of the paper, but here the more general case is considered.

The next step is to find the profit maximizing prices. Solving for the first-order condition w.r.t. price results in the following price for merchant 1:

$$p_1 = \frac{1}{2}[t + p_2 + \gamma + (1 - H(b_B^\star))(m - b_S)] \tag{36}$$

Since merchant 2 solves a symmetrical problem, the expression for p_2 is identical, only with subscripts interchanged. One can therefore solve for p_1 and p_2

$$p_1 = p_2 = t + \gamma + (1 - H(b_B^\star))(m - b_S) \tag{37}$$

Inserting the price expressions into the profit function produces

$$\pi_1 = \pi_2 = \frac{1}{2}t - K \tag{38}$$

To ensure that both merchants accepting cards is an equilibrium, it must be the case that neither has an incentive to deviate and accept cash only. To establish the conditions under which this will be the case, suppose therefore that merchant 2 only accepts cash. Letting x_C and x_D denote the market shares among cash and card users respectively, merchant 2 will obtain market shares of

$$(1 - x_D) = \frac{1}{2} + \frac{p_1 - p_2}{2t} \tag{39}$$

$$(1 - x_C) = \frac{1}{2} + \frac{p_1 - p_2 - E[b_B | b_B^\star]}{2t} \tag{40}$$

where the market share among card users is found by comparing the consumers' cost of trading with merchant 1 and paying by card, which is $p_1 + tx_C - b_B$, with the cost of trading with merchant 2, which is $P_2 + (1 - x_C)t$, and finally integrating over the interval $[b_B^\star, b_{B,max}]$.

Profits are

$$\pi_1 = H(b_B^\star)\left(\frac{1}{2}+\frac{p_2-p_1}{2t}\right)(p_1-\gamma) \tag{41}$$
$$+ (1-H(b_B^\star))\left(\frac{1}{2}+\frac{p_2-p_1+E[b_B|b_B\geq b_B^\star]}{2t}\right)(p_1-\gamma-m+b_S)-K$$

$$\pi_2=\left(\frac{1}{2}+\frac{p_1-p_2+(1-H(b_B^\star))(m-b_S+E[b_B|b_B\geq b_B^\star])}{2t}\right)(p_2-\gamma) \tag{42}$$

The associated first-order conditions are

$$p_1=\frac{1}{2}\left(t+\gamma+p_2+(1-H(b_B^\star))(m-b_S+E[b_B|b_B\geq b_B^\star]\right) \tag{43}$$

$$p_2=\frac{1}{2}\left(t+\gamma+p_1+(1-H(b_B^\star))E[b_B|b_B\geq b_B^\star]\right) \tag{44}$$

Solving for prices

$$p_1=t+\gamma+\frac{1}{3}(1-H(b_B^\star))\left(2(m-b_S)+E[b_B|b_B\geq b_B^\star]\right) \tag{45}$$

$$p_2=t+\gamma+\frac{1}{3}(1-H(b_B^\star))\left(m-b_S-E[b_B|b_B\geq b_B^\star]\right) \tag{46}$$

This implies profits for merchant 2 of

$$\pi_2 = \left(\frac{1}{2}+\frac{\frac{1}{3}(1-H(b_B^\star))(m-b_S-E[b_B|b_B\geq b_B^\star])}{2t}\right) \tag{47}$$
$$\times \left(t+\frac{1}{3}(1-H(b_B^\star))(m-b_S-E[b_B|b_B\geq b_B^\star])\right)$$

Hence, the condition required for a card acceptance equilibrium is that the above is less than or equal to $\frac{1}{2}t-K$. Examining this inequality and solving for b_S gives

$$b_S\geq m-E[b_B|b_B\geq b_B^\star]+\frac{3t(1-\sqrt{1-\frac{2K}{t}}}{1-H(b_B^\star)} \tag{48}$$

Proof of lemma 2

To show the existence of equilibria, one can appeal to Brouwer's fixed point theorem, which requires a continuous function $f: S\rightarrow S$ on a non-empty, convex, compact subset of a finite-dimensional normed linear space. Continuity of f follows from the continuity of $G(b_S)$ and $H(b_B)$. The set $S=[b_{S,min},b_{S,max}-\epsilon]$ is clearly convex, closed and bounded. Being a subset of the real numbers, the set is therefore compact. Hence, the only remaining requirement of Brouwer's fixed point theorem is that $f(b_S)$ maps S into itself, and imposing the above condition ensures that.

Proof of proposition 1

With $k=0$ there is only one equilibrium in which cards are used. The fraction of individuals using payment cards is $1-H(0)$. Applying lemma 2, the fraction of merchants accepting cards is

$$1-G\left(m-E[b_B\mid b_B\geq 0]+\frac{3t\left(1-\sqrt{1-\frac{2K}{t}}\right)}{1-H(0)}\right) \tag{49}$$

The fraction of card payments is consequently

$$\mu(m) = (1 - H(0)) \left(1 - G\left(m - E[b_B \mid b_B \geq 0] + \frac{3t\left(1 - \sqrt{1 - \frac{2K}{t}}\right)}{1 - H(0)}\right)\right) \tag{50}$$

Since $G\left(m - E[b_B \mid b_B \geq 0] + \frac{3t\left(1-\sqrt{1-\frac{2K}{t}}\right)}{1-H(0)}\right)$ is increasing in m, it follows that card usage is decreasing in m.

Proof of lemma 3

The expression, which simply says that banks in optimum equate marginal revenues and marginal costs, follows from solving for the banks' first-order condition and rearranging terms.

The second-order condition requires that

$$2g(m^\star) - g'(m^\star)(m^\star - c_C + c_D) < 0 \tag{51}$$

Inserting the expression for $m^\star$, one can verify that an increasing hazard ratio $\frac{g(b_S)}{1-G(b_S)}$ or, equivalently, that a decreasing inverse hazard ratio is a (more than) sufficient condition for this inequality to be satisfied.

Proof of proposition 2

Differentiation of the social objective function w.r.t. m and rearranging slightly results in the following expression

$$-g(m - E[b_B|b_B \geq 0]) \times \left(\int_0^{b_{B,max}} b_B dH(b_B) + (1 - H(0))\,(m - E[b_B|b_B \geq 0] - c_C + c_D)\right) = 0 \tag{52}$$

Dividing by $1 - H(0)$ and $-g(m - E[b_B|b_B \geq 0])$ and solving for m gives

$$m = c_C - c_D \tag{53}$$

The second-order condition is automatically satisfied at this value of m, independent of any distributional assumptions.

Proof of proposition 3

This is an immediate consequence of comparing the expressions found in lemma 3 and Proposition 2.

Proof of corollary 1

If m is set such that $b_S > m - E[b_B|b_B \geq 0] + \frac{3t(1-\sqrt{1-\frac{2K}{t}}}{1-H(0)}$, no merchant accepts card and no card payments take place. For any merchant fee that satisfies $b_S \leq m - E[b_B|b_B \geq 0] + \frac{3t(1-\sqrt{1-\frac{2K}{t}}}{1-H(0)}$, all merchants accept cards, and the fraction of card payments is $1 - G(0)$, which is independent of the merchant fee. The merchant acceptance threshold thus determines an upper bound for the optimal merchant fee, which is also the profit maximizing fee for banks.

The value of the social objective function is unchanged as m decreases. It does so until it reaches the point at which banks' rationality constraint is no longer satisfied. That point determines the lower bound of the interval.

Proof of lemma 4

Assume that merchant 1 accepts cards and surcharges. If both merchants accept cards and surcharges, they earn profits of $\frac{1}{2}t - K$. This can be shown using minor modifications to the proof in the appendix, in particular setting the adoption threshold at $B^\star = m$, removing the $-m$-term from the merchant profit functions, and then redoing the derivations. We must therefore establish the conditions under which merchant 2 is no better off by not accepting cards.

With surcharging, the fraction of consumers who always pay using cash when they can is $G(m)$. Among these, merchant 2 enjoys a market share, $1 - x_D$ of:

$$1 - x_D = \frac{1}{2} + \frac{p_1 - p_2}{2t} \tag{54}$$

Consumers in the remaining part of the population are indifferent between the two merchants whenever

$$p_1 - b_B + m + tx_C = p_2 + t(1 - x_D) \tag{55}$$

Integrating over all consumers with benefits in the interval $[m, b_{B,max}]$ gives a market share among card users (i.e. those who use cards when possible) of

$$1 - x_C = \frac{1}{2} + \frac{p_1 - p_2 - E[b_B|b_B \geq m] + m}{2t} \tag{56}$$

This implies profit functions of

$$\begin{aligned} \pi_1 &= H(m)\left(\frac{1}{2} + \frac{p_2 - p_1}{2t}\right)(p_1 - \gamma) \\ &+ (1 - H(m))\left(\frac{1}{2} + \frac{p_2 - p_1 + E[b_B|b_B \geq m] - m}{2t}\right)(p_1 - \gamma + b_S) - K \end{aligned} \tag{57}$$

$$\pi_2 = \left(\frac{1}{2} + \frac{p_1 - p_2 + (1 - H(m))(-E[b_B|b_B \geq m] + m)}{2t}\right)(p_2 - \gamma) \tag{58}$$

Deriving the two merchants' first-order conditions and solving for p_1 and p_2 gives

$$p_1 = t + \gamma + (1 - H(m))\left(-\frac{2}{3}b_S + \frac{1}{3}E[b_B|b_B \geq m] - \frac{1}{3}m\right) \tag{59}$$

$$p_2 = t + \gamma + (1 - H(m))(-\frac{1}{3}b_S - \frac{1}{3}E[b_B|b_B \geq m] + \frac{1}{3}m) \tag{60}$$

Inserting these expressions into the second merchant's profit function and simplifying gives

$$\pi_2 = \frac{1}{2t}\left[t + (1 - H(m))(-\frac{1}{3}b_S - \frac{1}{3}E[b_B|b_B \geq m] + \frac{1}{3}m)\right]^2 \tag{61}$$

Comparing this to the profit when both merchants accept cards and surcharge, one can show the no deviation condition to be

$$b_S \geq m - E[b_B|b_B \geq m] + \frac{3t\left(1 - \sqrt{1 - \frac{2K}{t}}\right)}{1 - H(m)} \tag{62}$$

6.1 Proof of proposition 4

For a given merchant fee, the fraction of payments made by cards is then greater under no surcharging if the following condition is satisfied

$$(1 - H(0))(1 - G(m - E[b_B|b_B \geq 0])) \geq (1 - H(m))(1 - G(m - E[b_B|b_B \geq m])) \tag{63}$$

Adding and subtraction terms produces

$$\begin{aligned} & (1 - H(0))(1 - G(m - E[b_B|b_B \geq 0])) \geq [(1 - H(0)) + (H(0) - H(m))] \\ \times \quad & [(1 - G(m - E[b_B|b_B \geq 0])) + (G(m - \delta) - G(m - E[b_B|b_B \geq m]))] \end{aligned} \tag{64}$$

Simplifying and rearranging

$$\frac{H(m) - H(0)}{1 - H(m)} \geq \frac{G(m - E[b_B|b_B \geq 0]) - G(m - E[b_B|b_B \geq m])}{1 - G(m - E[b_B|b_B \geq 0])} \tag{65}$$

When $m > 0$, $H(m) - H(0)$ is positive, while $E[b_B|b_B \geq m] > E[b_B|b_B \geq 0]$). Therefore $m - E[b_B|b_B \geq 0] > m - E[b_B|b_B \geq 0]$, and $G(m - E[b_B|b_B \geq 0]) - G(m - E[b_B|b_B \geq m]) < 0$. This means that inequality must be satisfied.

6.2 Proof of lemma 5

The expression is found by solving the banks' profit-maximization problem subject to the new rationality constraint.

Existence of equilibria with surcharging

In the following I consider the circumstances under which either surcharging or no surcharging are equilibria. To establish this, consider the situation in which merchant 1 does not surcharge. For no surcharging to be an equilibrium, merchant 2 must be no better off by surcharging.

In this case there are three distinct classes of consumers for whom market shares must be calculated for each merchant. Assuming that $m > 0$, for consumers with a strict preference for cash, i.e. $b_B < 0$, merchant 1's market share is:

$$x = \frac{1}{2} + \frac{p_2 - p_1}{2t} \tag{66}$$

Then there are consumers who will by card if transacting with merchant 1, but not with merchant 2. These are the ones who enjoy card benefits in the interval $[0, m)$. These consumers are indifferent between the two merchants when

$$p_1 + tx - b_B = p_2 + (1 - x)t \tag{67}$$

Solving for x, and integrating over the relevant interval, one finds that merchant 1 enjoys a market share of

$$\frac{1}{2} + \frac{p_2 - p_1 + E[b_B|0 \leq b_B < m]}{2t} \tag{68}$$

Finally, there are those consumers who will pay by card at either merchant, whether surcharging is applied or note. These enjoy card benefits $b_B \geq m$, and are indifferent between the two merchants at

$$p_1 + tx - b_B = p_2 + (1 - x)t + m - b_B \tag{69}$$

The calculations could also be done under the assumption that $m < 0$. In that case the three relevant intervals of consumers would be $[b_{B,min}, m)$ (cash users), $[m, 0)$ (cash users when transacting with merchant 1, and card users when transacting with merchant 2) and $[0, b_{B,max}$ (card users).

Among these types, merchant 1 therefore has a market share of

$$x = \frac{1}{2} + \frac{p_2 - p_1 + m}{2t} \tag{70}$$

Multiplying the market shares by profit margins for each class and weighing by the classes' proportion in the consumer population, one arrives at profit functions of

$$\begin{aligned}
\pi_1 &= H(0)(\frac{1}{2} + \frac{p_2 - p_1}{2t})(p_1 - \gamma) \\
&+ (H(m) - H(0))(\frac{1}{2} + \frac{p_2 - p_1 + E[b_B—0 \leq b_B < m]}{2t})(p_1 - \gamma + b_S - m) \\
&+ (1 - H(m))(\frac{1}{2} + \frac{p_2 - p_1 + m}{2t})(p_1 - \gamma + b_S - m) - K
\end{aligned} \tag{71}$$

$$\begin{aligned}
\pi_2 &= H(0)(\frac{1}{2} + \frac{p_1 - p_2}{2t})(p_2 - \gamma) \\
&+ (H(m) - H(0))(\frac{1}{2} + \frac{p_1 - p_2 - E[b_B—0 \leq b_B < m]}{2t})(p_2 - \gamma) \\
&+ (1 - H(m))(\frac{1}{2} + \frac{p_1 - p_2 - m}{2t})(p_2 - \gamma + b_S) - K
\end{aligned} \tag{72}$$

Deriving for the first-order conditions and solving for p_1 and p_2 gives

$$p_1 = t + \gamma + (H(m) - H(0))(-\frac{2}{3}b_S + \frac{2}{3}m + \frac{1}{3}E[b_B—0 \leq b_B < m]) + (1 - H(m))(-b_S + m) \tag{73}$$

$$p_2 = t + \gamma + (H(m) - H(0))(-\frac{1}{3}b_S + \frac{1}{3}m - \frac{1}{3}E[b_B—0 \leq b_B < m]) + (1 - H(m))(-b_S) \tag{74}$$

These expressions can be inserted into the profit function for merchant 2, the value of which must then be compared to $\frac{1}{2}t - K$. Doing this, one can find a closed-form expression which gives the values of of b_S, for which it does not pay off to deviate. However, the expression that results turns out to be rather involved, and not much intuition can be gleaned from it. The issue is examined numerically in section 5.3.

Using derivations that parallel the ones above, one can likewise find merchant profits and prices under the assumption of negative merchant fees. The resulting expressions are

$$\begin{aligned}
\pi_1 &= H(m)(\frac{1}{2} + \frac{p_2 - p_1}{2t})(p_1 - \gamma) \\
&+ (H(0) - H(m))(\frac{1}{2} + \frac{p_2 - p_1 - E[b_B—m \leq b_B < 0]}{2t})(p_1 - \gamma) \\
&+ (1 - H(0))(\frac{1}{2} + \frac{p_2 - p_1 + m}{2t})(p_1 - \gamma + b_S - m) - K
\end{aligned} \tag{75}$$

$$\begin{aligned}
\pi_2 &= H(m)(\frac{1}{2} + \frac{p_1 - p_2}{2t})(p_2 - \gamma) \\
&+ (H(0) - H(m))(\frac{1}{2} + \frac{p_1 - p_2 + E[b_B—m \leq b_B < 0]}{2t})(p_2 - \gamma + b_S) \\
&+ (1 - H(0))(\frac{1}{2} + \frac{p_1 - p_2 - m}{2t})(p_2 - \gamma + b_S) - K
\end{aligned} \tag{76}$$

with

$$p_1 = t + \gamma + (H(0) - H(m))(-\frac{1}{3}b_S - \frac{1}{3}E[b_B—0 \leq b_B < m]) + (1 - H(0))(-\frac{1}{3}b_S + \frac{1}{3}m) \quad (77)$$

$$p_2 = t + \gamma + (H(0) - H(m))(-\frac{2}{3}b_S + \frac{1}{3}E[b_B—0 \leq b_B < m]) + (1 - H(0))(-\frac{2}{3}b_S - \frac{1}{3}m) \quad (78)$$

References

Baxter, William (1983), "Bank interchange of transactional paper: Legal perspectives." *Journal of Law and Economics*, 26, 541–588.

Bolt, Wilko, Nicole Jonker, and Corry van Renselaar (2010), "Incentives at the counter: An empirical analysis of surcharging card payments and payment behaviour in the netherlands." *Journal of Banking and Finance*, 34, 1738–1744.

Börestam, Ann and Heiko Schmiedel (2011), "Interchange fees in card payments." *ECB Occasional Paper Series*, No. 131, 1–56.

Borzekowski, Ron, Elizabeth K. Kiser, and Shaista Ahmed (2008), "Consumers' use of debit cards: Patterns, preferences, and price response." *Journal of Money, Credit and Banking*, 40, 149–172.

Chakravorti, Sujit (2003), "Theory of credit card networks: A survey of the literature." *Review of Network Economics*, 2, 50–68.

Chakravorti, Sujit and Roberto Roson (2006), "Platform competition in two-sided markets: The case of payment networks." *Review of Network Economics*, 5, 118–142.

Danmarks Nationalbank (2011), "Costs of payments in Denmark."

DellaVigna, Stefano and Ulrike Malmendier (2006), "Paying not to go to the gym." *American Economic Review*, 96, 694–719.

Gans, Joshua S. and Stephen P. King (2003), "The neutrality of interchange fees in payment systems." *Topics in Economic Analysis and Policy*, 3, Article 1.

Guthrie, Graeme and Julian Wright (2007), "Competing payment schemes." *Journal of Industrial Economics*, 55, 37–67.

Heffernan, Shelagh A. (2002), "How do uk financial institutions really price their banking products?" *Journal of Banking and Finance*, 26, 1997–2016.

Jacobsen, Johan G. K. and Anders M. Pedersen (2012), "Cost of card and cash payments in denmark." *Danmarks Nationalbank Monetary Review*, 2nd Quarter, 109–121.

Jonker, Nicole (2011), "Card acceptance and surcharging: the role of costs and competition." *Review of Network Economics*, 10, Article 4.

Jonker, Nicole (2013), "Social costs of pos payments in the netherlands 2002-2012: efficiency gains from increased debit card usage." *DNB Occasional Studies*, 11.

McAndrews, James and Zhu Wang (2012), "The economics of two-sided payment card markets: Pricing, adoption and usage." *FRB Richmond Working Paper Series*, No. 12-06.

Rochet, Jean-Charles and Jean Tirole (2002), “Cooperation among competitors: Some economics of payment card associations.” *The RAND Journal of Economics*, 33, 549–570.

Rochet, Jean-Charles and Jean Tirole (2006), “Two-sided markets: A progress report.” *The RAND Journal of Economics*, 37, 645–667.

Rochet, Jean-Charles and Jean Tirole (2011), “Must-take cards: Merchant discounts and avoided costs.” *Journal of the European Economic Association*, 9, 462–495.

Rochet, Jean-Charles and Julian Wright (2010), “Credit card interchange fees.” *Journal of Banking and Finance*, 34, 1788–1797.

Rysman, Marc (2009), “The economics of two-sided markets.” *Journal of Economic Perspectives*, 23, 125–143.

Schmalensee, Richard (2002), “Payment systems and interchange fees.” *Journal of Industrial Economics*, 50, 103–122.

Schmiedel, Heiko, Gergana Kostova, and Wiebe Ruttenberg (2012), “The social and private costs of retail payment instruments.” *ECB Occasional Paper Series*, No. 137, 1–49.

Shy, Oz and Zhu Wang (2011), “Why do payment networks charge proportional fees.” *American Economic Review*, 101, 1575–1590.

Verdier, Marianne (2011), “Interchange fees in payment card systems: A survey of the literature.” *Journal of Economic Surveys*, 25, 273–297.

Vickers, John (2005), “Public policy and the invisible price: Competition law, regulation and the interchange fee.” *Competition Law Journal*, 4, 5–16.

Wang, Zhu (2012), “Debit card interchange fee: Some assessments and considerations.” *Federal Reserve Bank of Richmond Economic Quarterly*, 98, 159–183.

Wright, Julian (2003), “Optimal card payment systems.” *European Economic Review*, 47, 587–621.

Wright, Julian (2004), “The determinants of optimal interchange fees in payment systems.” *Journal of Industrial Economics*, 52, 1–26.

Wright, Julian (2012), “Why payment card fees are biased against retailers.” *The RAND Journal of Economics*, 43, 761–780.

Zenger, Hans (2011), “Perfect surcharging and the tourist test interchange fee.” *Journal of Banking and Finance*, 35, 2544–2546.

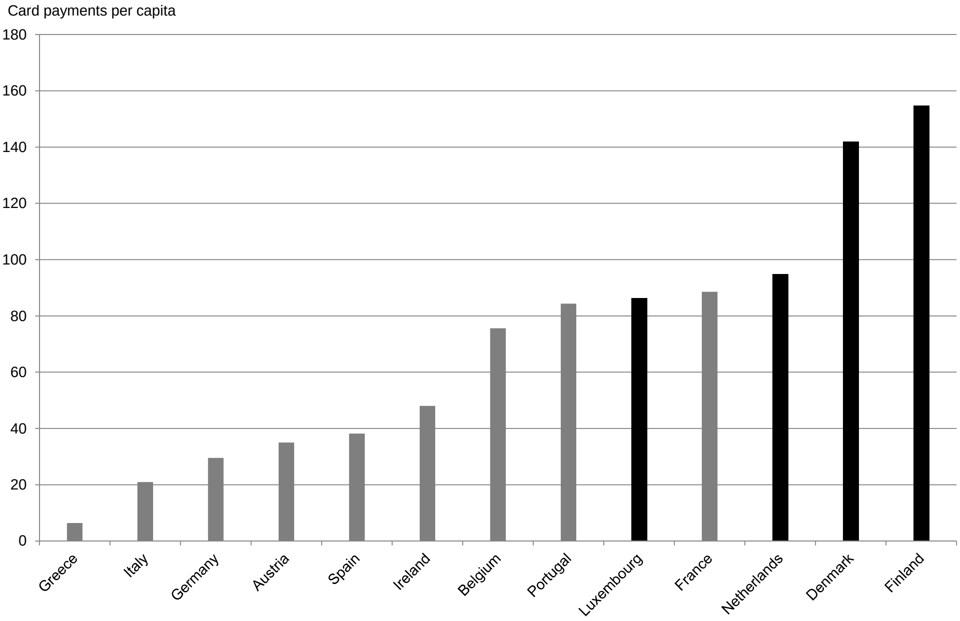

Figure 1: Card usage in selected European countries
The figure shows the number of card payments per capita per year in a number of European countries. Data source: ECB Statistical Data Warehouse.

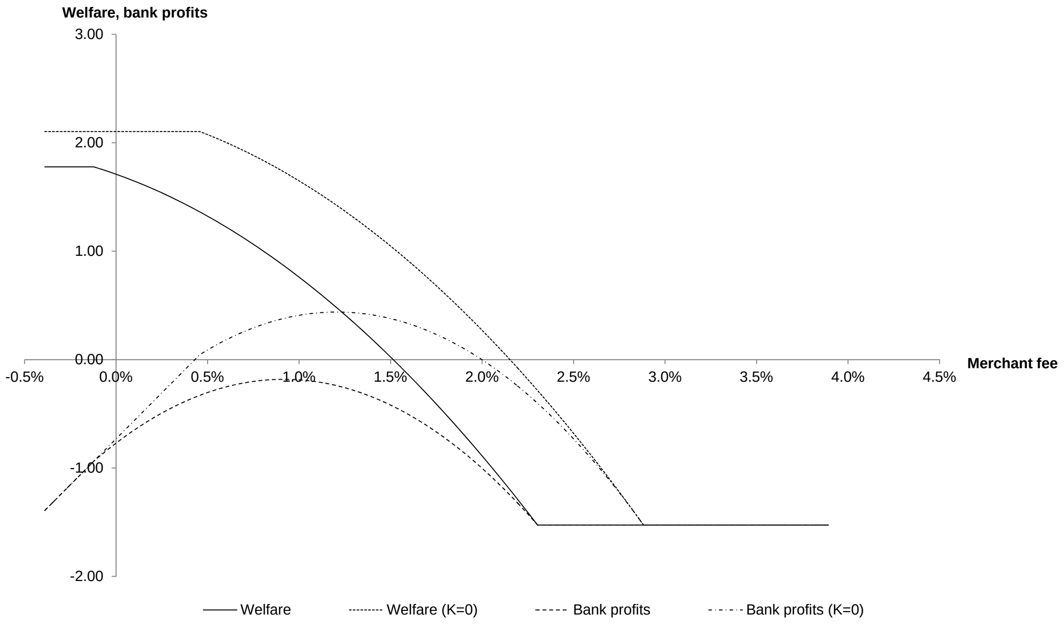

Figure 2: Welfare and profits (without surcharging)
The figure shows welfare and profits as a function of the merchant fee, under the assumption that merchants do not surcharge the fees. The upper lines represent welfare, the lower lines profits. Welfare and profits is shown under both the assumption of no merchant fixed costs and merchant fixed costs. The flat portions of the lines represent areas where either all (to the left) or no (to the right) consumers use cards.

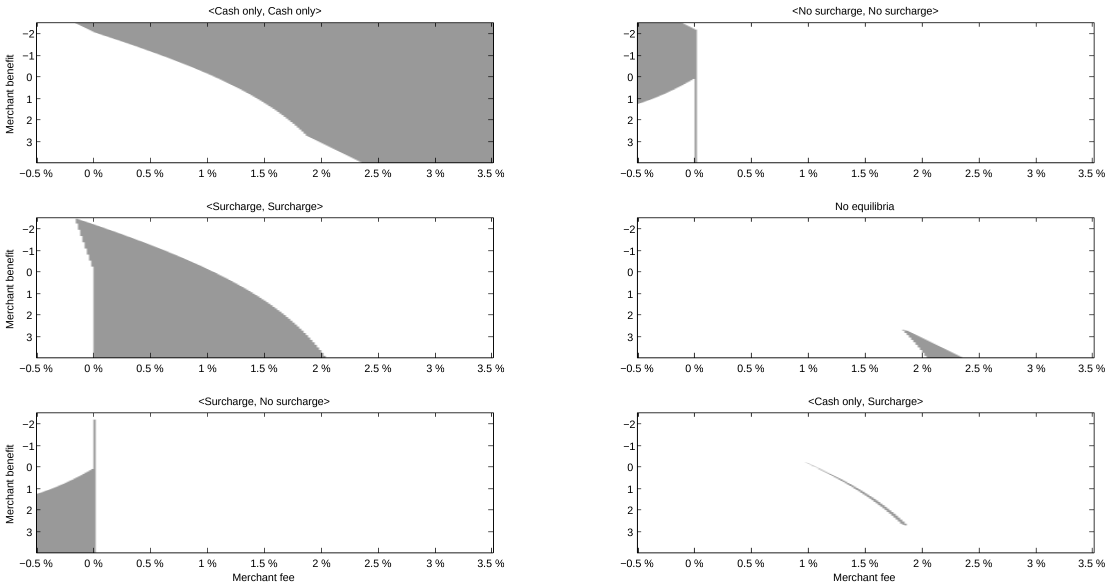

Figure 3: Equilibria in model with surcharging
The shaded area of each figure represents the combinations of merchant fees (x-axis) and merchant benefits (y-axis) for which the response profile in question (see figure titles) is a Nash equilibrium.

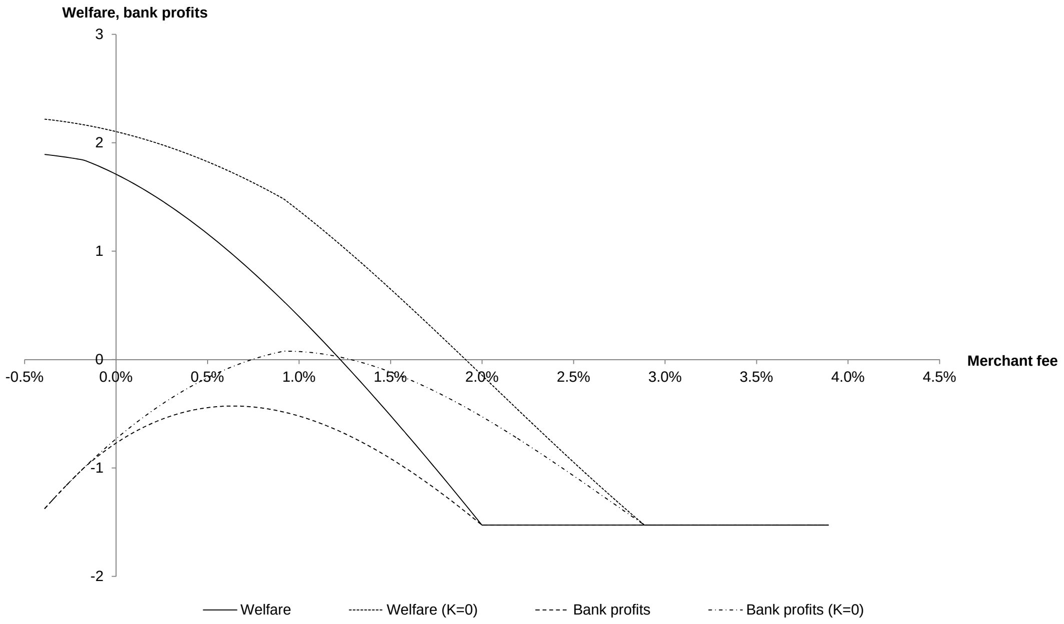

Figure 4: Welfare and profits (with surcharging)
The figure shows welfare and profits as a function of the merchant fee, under the assumption that merchants surcharge the fees. The upper lines represent welfare, the lower lines profits. Welfare and profits is shown under both the assumption of no merchant fixed costs and merchant fixed costs. The flat portions of the lines represent areas where either all (to the left) or no (to the right) consumers use cards.

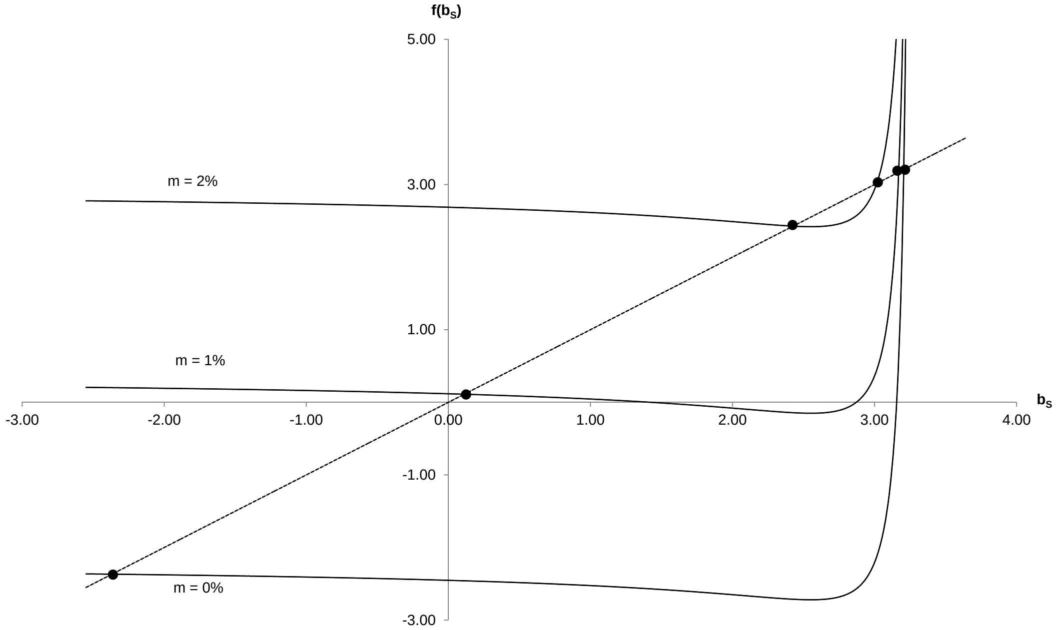

Figure 5: Fixed points with consumer and merchant fixed costs
The figure depicts, for three different merchant fees, values of the threshold of the merchant benefits for which the fraction of merchants accepting cards is consistent with the fraction of consumers using cards (and vice versa). There are two fixed points for each merchant fee, the leftmost corresponding to an equilibrium with high card usage and acceptance and the rightmost corresponding to an equilibrium with low card usage and acceptance.

Table 1: Estimates of model parameters

Parameter	Estimate
γ	230
t	27
K	0.326
c_C	0.321
c_D	1.527
$b_{S,max}$	3.698
$b_{S,min}$	-2.533
$b_{B,max}$	7.421
$b_{B,min}$	-3.823