Credit and the No-Surcharge Rule

Cyril Monnet
European Central Bank
cyril.monnet@ecb.int

William Roberds
Federal Reserve Bank of Atlanta
william.roberds@atl.frb.org

November 1, 2006

Abstract

A controversial aspect of payment cards has been the "no-surcharge rule." This rule, which is part of the contract between the card provider and a merchant, states that the merchant cannot charge a customer who pays by card more than a customer who pays by cash. In this paper we consider the design of an optimal card-based payment system when cash is available as an alternative means of payment. We find that a version of the no-surcharge rule emerges as a natural and advantageous feature of such a system.

1 Introduction

Even the most casual observer will have noticed that money is being displaced by memory. Transactions that were once only conducted only with cash–purchases of groceries, fast food meals, movies, taxi rides, etc.–are increasingly made using credit cards, debit cards, and various other payment methods that electronically link buyers to their payment histories.[1]

A commonly cited explanation for this phenomenon is the ongoing improvement in information technology. As the cost of storing, transmitting, and authenticating data falls, so the thinking goes, payment systems based on electronic accounts become more attractive relative to cash. This explanation is in accordance with accepted monetary theory (Kocherlakota 1998), which views cash as a second-best proxy for credit when the latter is too costly or simply unavailable.

A practical difficulty with this argument, however, is that in many instances the cost of making a cash transaction remains noticeably lower than any other type of payment. The cost of simply handing over a banknote, after all, is still virtually zero, and the burden imposed by inflation has fallen drastically over the past two decades. Systematic studies, taking into account the costs of safekeeping, trips to the bank, etc. place the merchant's cost of a typical cash transaction in the U.S. at about \$.10. By contrast, the average merchant cost of a debit card transaction is in the range of \$.34, and the typical credit card transaction costs a merchant in excess of \$.70.[2] A recent study by Garcia-Swartz, Hahn, and Layne-Farrar (2006) attempts to measure the costs of various payment methods to all parties involved. It argues that card payments are more competitive with cash, once buyers' "implicit cost" of using cash–

[1] Aggregate statistics collected by the Bank for International Settlements (Committee on Payment and Settlement Systems, 2006) show that the volume and value of card-based payments has sharply accelerated over the past ten years in all developed countries. This trend is also apparent in U.S. household survey data (Klee 2006) and recent Federal Reserve surveys (Gerdes at al. 2005). While these numbers do not track cash payments, a 2005 survey conducted by Visa, cited in Garcia-Swartz, Hahn, and Layne-Farrar (2006) indicates a dropoff in the use of cash in the U.S. over the past decade.

[2] Figures are from an oft-cited 2001 Food Marketing Institute survey; see Humphrey et al. (2003).

especially the "shoe-leather" cost of visiting an ATM, estimated at more than \$.28 per transaction–is taken into account. As this last number is necessarily somewhat imprecise (for example cash can circulate in ways that are not easily observable by econometricians), we would still argue that cash remains the cheapest way to pay in many situations.

The dominant component of the cost of a card payment is the *merchant fee* (a.k.a. "merchant discount") paid by a seller of goods or services to the card company. In the U.S. this fee averages about 2% of the purchase amount.[3] Usually this fee is not paid explicitly by the buyer, but is instead deducted from the merchant's payment by the card provider, and the buyer pays the same price as he would have using cash. Payment card providers reinforce this practice with a contractual provision known as a *no-surcharge rule* (NSR, a.k.a. "no-discrimination rule") that prohibits merchants from assessing a fee on customers who wish to pay with their credit or debit card.[4]

The no-surcharge rule has been extremely controversial, and has been banned in some countries (e.g., Australia; see Lowe 2005) as a form of collusive price-fixing. Critics of NSR have argued that it inefficiently encourages the use of more costly forms of payment (credit cards) to less costly (cash), leading to what the Governor of the Reserve Bank of Australia has termed a "Gresham's Law of Payments" (Macfarlane 2005).[5]

Against the cost disadvantages must be set the benefits of card payment: certainly in the case of credit cards at least, paying by card allows buyers to tap into their credit lines in a convenient and straightforward way. But this argument does not apply in the case of debit cards or credit cards that are paid off every month. "Paybacks"

[3] Average merchant fess on credit card transactions in 2005 were 2.19% for MasterCard and Visa, 2.41% for American Express, and 1.76% for Discover (*Nilson Report,* Issue 862, 2006).

[4] Under U.S. law consumers are still entitled to negotiate discounts if they offer cash. Such discounts are rarely offered for routine purchases, however. In some other countries even this practice is prohibited.

[5] See Chakravorti and To (1999) for a formal presentation of this idea. Lowe (2005) notes that surcharging for card use is still uncommon in Australia, despite the regulatory removal of NSR.

to card use, in the form of frequent flyer miles, cash-back, or other rewards, further increase buyers incentives to use cards, but paybacks do not explain why still-cheaper cash is not preferred.

Of course, the Coase Theorem would predict that the NSR is irrelevant, as long as all parties to a transaction are able to contract around it. Papers in the industrial organization literature, such as Rochet and Tirole (2002), contend that the Coase Theorem can fail in a payments environment, due to an asymmetry in market power between merchants and consumers.[6] Absent a no-surcharge rule, it is argued, monopolistic merchants may inefficiently shift the costs of a card payment system to consumers, leading to thet underprovision of credit and welfare losses.

Below, we abstract from industrial organization issues and instead focus on other frictions that could cause the NSR to matter and the Coase Theorem to fail. These are the standard frictions that give rise to the use of payment systems: time mismatches of agents' trading demands, private information about agents' preferences, and limited enforcement of their pledges to repay. As in actual payment situations, the term "limited enforcement" incorporates both the potential anonymity of buyers and sellers, and, once identification has occurred, a limited ability to apply penalties when an agent defaults.

Next we introduce a transactions technology, which, at a cost, allows for relaxation of these frictions. This technology, which we interpret as a credit-card payment system, must compete with an alternative payments technology in the form of cash. To make cash as attractive as possible, we assume it is uncounterfeitable, not subject to theft, can transferred for free, and that it bears little or no inflation tax. We then consider how a planner would structure an optimal credit card system when cash is available, and find that a version of the no-surcharge rule can emerge as an advantageous feature of such a system.

The intuition behind this result is simple. In our environment, agents early in

[6] For surveys of the extensive I/O literature on card payments see Chakravorti (2003), Hunt (2003), Rochet and Tirole (2004), Evans and Schmalensee (2005), and Rochet and Tirole (2006).

their life-cycle always need credit, whereas older agents never do. For young agents to access credit, they must forsake the anonymity associated with cash payment and join the credit-card payment system. But young agents' participation in this system is pointless, unless a sufficiently large number of older agents also agree to take part. A no-surcharge rule promotes ongoing participation in credit arrangements by, in effect, taxing the use of cash. Thus, in our model, the use of credit card system has both a private and social benefit, and a no-surcharge rule allows agents to internalize the effects of their participation.

This result should be cast in a proper light. Since our analysis pertains to the design of an ideal card payment system, it renders no verdict on pricing structures in existing card payment systems. Rather, it shows how the oft-discussed "network benefits" of card payments can arise in a general-equilibrium world, and how an absence of surcharges can be instrumental in the capture of these benefits.

2 The Environment

Our basic environment closely follows Berentsen, Camera and Waller (forthcoming) which in turn derives from Lagos and Wright (2005).

Agents, Preferences, and technology. Time $t = 0, 1, ...$, is infinite and discrete. The economy consists of a $[0, 1]$ continuum of agents. A small group (measure zero) of these agents live forever and discount the future at rate δ. All other agents have finite lives, and die each period with probability $1 - \delta$. A group of agents with measure $1 - \delta$ is also born during each period, so that the size of the population does not change from period to period. Associated with each agent is a unique fixed quantity that we will call his "identity."[7]

Agents produce and consume two types of nonstorable goods: a specialized good and a numeraire good. Utility (disutility) of consumption (production) of the nu-

[7] Alternatively, an agent's identity could be thought of as his "location," although our model does not rely on geographical dispersion.

meraire good is linear. Consumption of the numeraire good is denoted as the negative of "hours worked" h.

There is a linear cost $-q_s$ of producing q_s units of the specialized good. All agents enjoy utility $u(q)$ when they consume q units of the specialized good, where $u'(q) > 0$, $u''(q) < 0$, $u'(0) = \infty$ and $\lim_{q\to\infty} u'(q) = 0$.

Trading stages. Each period has two sub-periods, each with its own market. Markets in both sub-periods are Walrasian in the sense that no agent has market power. The nature of trading is quite different in the two subperiods, however.

In the first subperiod, the numeraire good is exchanged and a trader's identity can be verified by other agents at zero cost, should the agent decide to make his identity available. We call this stage the *settlement stage*. At the end of the settlement stage a randomly selected proportion $1-\delta$ of the finite-lived agents die and are replaced by young agents. Upon reaching their second settlement period, finite-lived agents become inherently indistinguishable from other agents and are referred to as "adults."

In the second sub-period, the *trading stage*, the specialized good is exchanged in "incomplete anonymity." This means that in this stage, agents are unrecognizable to one another, absent the application of some costly identifying technology (discussed in detail below). Application of this technology allows an agent's identity to be determined with perfect accuracy; however agents again always decide whether they want to be identified and may remain anonymous if they so prefer.

Within the trading stage, a proportion $a \in (0,1)$ of the population has the opportunity to actively engage in trade. Of these, a randomly selected proportion $1-n$ have the desire to consume specialized goods produced by others; otherwise they become potential producers with probability $n \in (0,1)$. The remaining proportion $1-a$ of the population are inactive, meaning that during this stage they have neither the desire to consume nor the opportunity to produce. A consumer's trading-stage state (buyer, seller, inactive) is private information. The timing of events within a period is summarized in the following table.

Table 1: Events within a period

1. Settlement stage
 1a. Agents produce, trade, and consume numeraire good
 1b. $1-\delta$ adult agents die and are replaced by young agents

2. Trading stage
 2a. Agents learn if they are consumers, producers, or inactive
 2b. Buyers and sellers trade specialized goods
 2c. Buyers consume specialized goods

Immortal agents have exactly the same utility, endowments, and trading opportunities as finite-lived agents, differing in the detail that they do not die. The population of immortal agents remains constant at some low (zero measure) level.

Allocations. In this economy, we will consider symmetric stationary feasible allocations. A stationary allocation is a vector $(q, q_s, q^y, q_s^y, h_s, h_b, h_i)$, where the superscript y denotes young agents (superscripts are omitted for adults). (h_s, h_b, h_i) are hours worked for an agent of type j when he was a seller (s), buyer (b), or inactive (i) in the previous trading stage. The allocation is feasible if

$$
\begin{aligned}
(1-\delta)[(1-n)\, q^y - n q_s^y] + \delta[(1-n)\, q - n q_s] &= 0, \\
a\,(n h_s + (1-n)\, h_b) + (1-a)\, h_i &= 0.
\end{aligned}
$$

The first equality is the market clearing condition for the specialized good in the trading stage. The second constraint is the market clearing condition for the numeraire good during the settlement stage.

2.1 First best allocation

We begin by considering optimal allocations in the absence of information and enforcement constraints. The expected utility of a representative young agent just born, in his very first period of life, is

$$W^y = a\{(1-n)[u(q^y) - h_b] - n(q_s^y - h_s)\} - (1-a)h_i.$$

The expected utility of an adult at the start of a trading stage is likewise

$$W = a\{(1-n)[u(q) - h_b] - n(q_s - h_s)\} - (1-a)h_i.$$

We take the planner's objective function to be the maximization of the population-weighted sum of the agents' utility (ignoring the zero-measure group of immortal agents)

$$\mathcal{W} = (1-\delta)\, W^y + \delta W$$

Using the market clearing conditions, this reduces to

$$\mathcal{W} = a\{(1-\delta)(1-n)\,[u(q^y) - q^y] - \delta(1-n)\,[u(q) - q]\}$$

Therefore a first best allocation is one that satisfies $q^y = q = q^*$ where

$$u'(q^*) = 1.$$

We will denote the first best allocation by q^*. In words, a first-best allocation requires that the utility of consuming an additional increment of the specialized good equals the cost of producing this increment, for almost *all* consumers, i.e., for both the young and adults.

3 The cash economy

As agents are not readily indentifiable during the trading stage, they need some type of recordkeeping in order to transact. In this section we consider the economy where only cash – the simplest form of recordkeeping – is available. Cash is uncounterfeitable and can be costlessly authenticated and transferred among agents. Agents may trade numeraire for cash during the settlement stage; cash so acquired can then be used for purchases during the trading stage.

Cash is provided by a central bank. Let M be the per capita supply of cash. A transfer of cash takes place at the beginning of the settlement stage. More precisely the central bank makes a lump sum transfer of τM to all (adult) agents in the settlement stage. A share $(1-\delta)$ of this transfer reverts to the central bank when some of the agents die. The central bank then redistributes it pro-rata to alive adults. Hence, alive adults get a transfer of $\tau' M = [\tau + (1-\delta)\tau/\delta]\, M$. The net stock of money however grows as $M_{+1} = (1+\tau)M$. We will concentrate the analysis on stationary monetary equilibria where $\phi M = \phi_{-1} M_{-1}$, ϕ being the real price of money in terms of numeraire. We will denote the growth rate of money M/M_{-1} by γ. Hence γ also equals ϕ_{-1}/ϕ and we have

$$\gamma = 1 + \tau. \tag{1}$$

Let $V(m)$ denote the discounted lifetime utility of an adult when he enters the settlement stage holding m units of cash, while $W(m)$ denotes the expected discounted lifetime utility from entering the trading stage with money holding m. $V(m)$ is defined as

$$\begin{aligned} V(m) &= \max_{h, m_{+1}} -h + \delta W(m_{+1}) \\ & \quad\; s.t. \quad \phi m_{+1} = h + \phi m + \phi \tau' M. \end{aligned}$$

The first-order and envelope conditions give

$$\delta W'(m_{+1}) = \phi, \qquad V'(m) = \phi. \tag{2}$$

It follows that m_{+1} is independent of the past trading history of agent which is summarized in m, and all adults exit the settlement stage with the same holdings of cash.

The discounted lifetime utility of *adults* when they enter the trading stage with m units of cash is

$$W(m) = a\{(1-n)[u(q) + V(m - pq)] + n[-q_s + V(m + pq_s)]\} + (1-a)V(m),$$

where q and q_s are set optimally as follows.

An adult *producer* at the trading stage solves

$$\max_{q_s} -q_s + V(m + pq_s),$$

with first-order condition

$$pV'(m + pq_s) = 1.$$

Therefore using (2) we have

$$p\phi = 1. \tag{3}$$

The problem of an adult *consumer* is

$$\begin{aligned} \max_{q} \quad & u(q) + V(m - pq) \\ & s.t. \quad pq \leq m. \end{aligned}$$

and the first order condition gives

$$u'(q) - \phi p = \lambda p,$$

where $V'(m - pq)$ has been replaced by ϕ using (2). That is to say, either $\lambda > 0$ in which case the consumer's budget constraint binds so that $q = m/p$ and $u'(q) > \phi p$, or the budget constraint does not bind, $\lambda = 0$ and q solves $u'(q) = p\phi$.

The discounted lifetime utility of a *young agent* entering the trading stage is given by

$$\begin{aligned} W^y(m) \quad = \quad & a\{(1-n)[u(q^y) + V(m - pq^y)] + n[-q_s^y + V(m + pq_s^y)]\} \\ & +(1-a)V(m), \end{aligned}$$

where $m = 0$. The problem of a young producer and a young consumer are identical to their adult counterparts. The solution of a young producer's problem is identical to that of an adult; however a young consumer must choose $q^y = 0$ as he lacks the necessary cash to make a purchase.

The market clearing condition for the trading stage reflects the fact that young would-be buyers cannot consume, and is given by

$$\delta \left(\frac{1-n}{n} \right) q = q_s = q_s^y.$$

It is now easy to determine the value of an additional unit of money when (adult) agents exit the settlement stage. Using the solution to the buyer's problem, this is

$$W'(m) = a(1-n)[(u'(q)-1)\frac{1}{p} + \phi] + [an + (1-a)]\,\phi.$$

However, we know that $p\phi = 1$ and $W'(m) = \phi_{-1}/\delta$. Hence we obtain another equilibrium condition, which characterizes the quantity consumed in the trading stage as a function of the money growth rate γ

$$\frac{\gamma - \delta}{\delta} = a(1-n)[u'(q) - 1]. \tag{4}$$

We can now define and characterize a stationary monetary equilibrium.

Definition 1 *A stationary monetary equilibrium is a list (γ, q) such that (1) and (4) hold.*

Proposition 1 *A stationary monetary equilibrium exists for all $\gamma \geq \delta$ and is unique for all $\gamma > \delta$. The equilibrium allocation q is strictly decreasing in γ.*

The proof of the proposition follows immediately from condition (4). Note that if the Friedman rule holds, $\gamma = \delta$ and adult consumers always hold enough money to purchase the efficient amount, i.e., in equilibrium $u'(q_m) = 1$. This does not attain the first best however, as young buyers must postpone consumption until they have acquired the necessary stock of cash.

4 The cash-credit economy

When all trading-period transactions are in cash, young agents cannot consume in their first period of life. The requirement to transact in cash also constrains old agents when monetary policy does not follow the Friedman rule. Hence the availability of credit can potentially improve welfare by relaxing agents' cash constraints. In this section, we consider how a credit arrangment affects welfare. This part of the paper draws on the private information approach used by Koeppl, Monnet, and Temzelides (2006, KMT).

The anonymity prevalent during the trading periods means that agents cannot increase their consumption by simply issuing bonds: absent some means of identifying the bond issuer, such a bond would be worthless. Consequently the credit arrangement must also incorporate some technology for identifying debtors. To fix ideas, we imagine that this identification occurs using credit cards.

More specifically, we may imagine that the planner relies on a club arrangement known as a *credit card institution* (CCI). The CCI incurs monitoring and set-up costs $\mu \geq 0$ in ascertaining the identity of new credit card holders ("members").[8] This cost is denominated in the numeraire good and must be borne by agents using the CCI. Once it has incurred this cost, the CCI can costlessly verify agents as members and record a member's transactions. However it cannot observe whether an agent's state is "active" or "inactive" at the trading stage, or whether an agent trades with cash in the trading stage. The identities of infinite-lived agents are unverifiable, which prevents them from joining the CCI.

The CCI will seek to implement the first best allocation q^* in the trading stage. Hence, CCI does not seek to maximize profit but only to achieve q^* and to recover

[8] As in Kahn and Roberds (2005, KR), each new member, upon verification of his identity, receives a unique, uncounterfeitable credit card that may be subsequently verified at zero cost. Following KR, we will exclude the possibility of fraud on the supplier ("merchant") side. This means that in every trading period, once a CCI member identifies himself as a supplier, his delivery (or nondelivery) of specialized goods *within that period* becomes observable.

costs. Following KMT, we will design the terms of the credit card institution so that agents truthfully reveal their state (consumer, producer, inactive) and so that they are willing to participate in the credit card arrangement given the outside option of using cash. The CCI has no control over monetary policy and takes the money growth rate γ as given.

As in KMT, we assume that the CCI assigns credit *balances* to participants. It specifies rules for how the balances are updated given the histories of reports regarding transactions in the trading round. During the settlement stage, participants can trade balances for the numeraire good. Here the settlement stage is modelled as a competitive market in which agents who are "low" can increase their balances by producing numeraire, while those with high balances end up as consumers of numeraire. During the settlement stage, balances are exchanged for numeraire at price ϕ_b. We let d denote the amount of balances with which agents exit the settlement stage.

Momentarily ignoring agents' balances, there are three possibilities regarding meetings: an agent can be a consumer, a producer, or inactive. The vector of policy rules (L_t, K_t, B_t, q^*) then determines the respective balance adjustments for each type of agent and the quantity consumed by each consumer, q^*. More precisely, $L_t(K_t)$ is the adjustment for an agent who consumes (produces), while B_t is the adjustment for an inactive agent. These functions in general may depend on the agents' histories of transactions, as summarized by their current balances, as well as on the distribution of balances when agents exit the settlement stage. Balances are represented by real numbers not restricted in sign, while production of goods in the trading stage is restricted to be positive. After each trading stage, agents enter the settlement round knowing their new balances.

We should note that the CCI is subject cannot impose any direct penalty on a member with low balances who does not readjust his balance during a settlement period. That is, a CCI member can walk away from the arrangement at any time.

The only penalty that the CCI can apply is denial of future access.[9]

We define a *credit card system* (CCS) to be an array of functions $\{L_t, K_t, B_t, q^*\}$. A CCS is *feasible* if it satisfies some incentive and participation constraints (specified below). A CCS is *simple* if balance adjustments do not depend on the agents' current balances and are therefore history independent. A feasible CCS is optimal if it implements the efficient trading-period allocation q^*. Note that a CCS requires the existence of a CCI in order to identify agents.

Assumption 1 *The first best allocation satisfies* $\delta u(q^*) > q^*$.

Under assumption 1, KMT show that there exists a simple optimal CCS, where balances upon exiting the settlement stage d_t are equal to zero for all t. In other words, if a member of the CCI exits the settlement stage with $d \neq 0$, then the CCI shuts down the account of this member so that his credit card becomes invalid. It follows that the equilibrium distribution of balances is degenerate at $d = 0$, when CCI members exit the settlement stage.

As in the cash economy, the trading stage is still a Walrasian market. The auctioneer has a list transmitted by the CCI of those eligible to use credit. The auctioneer then calls a price p and quantities $(q_m, q_{s,m})$ consumed and produced when cash is used, as well as the terms of the CCS. Then CCI members decide if they will participate, and if so, whether they will use cash or credit. That is, they may participate anonymously as cash agents, or they may allow themselves to be identified, and participate in the credit market. CCI members thus have the opportunity to use cash rather than credit at any time. Infinite-lived agents remain outside the CCI, and will always use cash. Hence, while an active cash market will always exist, a successful CCS should contain incentives such that CCI members transact using credit.

In such a system, prices and quantities have to clear the market. We assume that the auctioneer does not cross-subsidize consumption across those agents that use cash

[9]This severe limitation on enforcement is largely consistent with industry experience. For example, in 2005, MasterCard and Visa were only able to collect about 9 percent of defaulted credit card balances (*Nilson Report* Issue 851, 2006).

and those that use credit.[10] More precisely, the auctioneer faces two market clearing conditions. Given that active agents are sellers with probability n and buyers with probability $1-n$, these market clearing conditions are

$$\begin{aligned}(1-n)\,q^* &= nq_s^*,\\ (1-n)\,q_m &= nq_{s,m}.\end{aligned}$$

We also exclude the possibility of cross-subsidization through the central bank: only agents holding currency in the settlement period are eligible to receive lump-sum transfers from the central bank. This can occur, for example, if currency becomes worn after a single period, so that cash holders must submit old banknotes in order to obtain new ones.

In the previous section, we have studied the problem of agents who only have access to cash. We now describe the problem of agents with access to the CCI.

4.1 Settlement stage

Let $Z(b,m)$ denote the value function of a CCI member who exits the trading round with balance b and cash holdings m. Let $H(d,m_{+1})$ denote the value of an agent who exits the settlement round with balance $d=0$ and cash holdings m_{+1}. Given the price in the settlement round is ϕ_b, CCI members at the beginning of the settlement round solve the following:

$$\begin{aligned} Z(b,m) &= \max_{h,m_{+1}}\{-h+\delta H(b_{+1},m_{+1})\} \qquad (5)\\ \text{s.t.}\quad \phi m_{+1}+\phi_b b_{+1} &= h+\phi_b b+\phi m.\end{aligned}$$

The first order condition with respect to money gives

$$\delta H_m\,(0,m_{+1})\le\phi \text{ with strict inequality if } m_{+1}=0. \qquad (6)$$

[10] Some policy discussions of NSR have focused on the issue of potential cross-subsidization between purchasers using cash and those using cards. The present model does not address this issue since, in equilibrium, the CCS will be utilized by (almost) all agents or none. We abstract from this issue in order to focus on the patterns of "subsidization" across heterogeneous cardholders, that are needed to sustain the CCS when cash is available as an alternative.

The envelope conditions are $Z_b = \phi_b$ and $Z_m = \phi$. Since the CCI requires credit users to carry zero balances when they exit the settlement stage, under the threat of credit exclusion, we have $b_{+1} = 0$. When CCI members do not carry cash $m_{+1} = 0$ and it follows that

$$h = -\phi_b b.$$

Note that the linearity in preferences implies that the value function $Z(b, m)$ is linear in credit balances and cash holdings.

Also, for any b, it should be the case that CCI members are better off (at the settlement stage) staying in the system than choosing to use cash forever after. This *no-default constraint* imposes that for any b,

$$Z(b, m) \geq \phi(m - m_{+1}) + \delta W(m_{+1}).$$

In the Appendix, we show that this requirement reduces to

$$\phi_b L - \delta\mu \geq -\gamma q_m + \frac{\delta a(1-n)}{1-\delta}\left[u(q_m) - u(q^*) - (q_m - q^*)\right] \tag{7}$$

The *cost-recovery constraint* in the settlement stage requires that CCI members cover monitoring costs for young agents joining the CCI[11]

$$a\phi_b\left[nK + (1-n)L\right] + (1-a)\phi_b B + (1-\delta)\mu = 0. \tag{8}$$

4.2 Trading stage

We now turn to the problem faced by CCI members during the trading round. Agents make reports to the auctioneer about their state (consumer, producer, inactive). Those that report "producer" receive instructions from the auctioneer to produce

[11] Expression (8) is somewhat restrictive in that it apportions the verification costs equally across all CCI members. More complicated schemes are feasible, but these are not observed in practice and are unlikely to change the results below. In particular, note that the planner cannot simply assign the verification costs to the beneficiaries of the CCI–the young. For sufficiently low inflation rates, this would imply that every young consumer would default during his first settlement stage and the CCI would collapse.

q_s^*. Consumers receive q^*. The auctioneer subsequently communicates the identity and reports of those CCI members that made use of the CCS to the CCI, which then makes balance adjustments depending on these reports.

For agents to report their state truthfully, some conditions have to be satisfied. In particular, *incentive constraints* require that the following inequalities hold:

$$\begin{aligned} u\left(q^*\right) + Z(L,m) &\geq Z(B,m), \\ -q_s^* + Z(K,m) &\geq Z(B,m), \\ Z\left(B,m\right) &\geq Z\left(L,m\right). \end{aligned}$$

The first (second) constraint states that a consumer (producer) must be at least as well off declaring his true state than reporting he is inactive. The third constraint states that an inactive agent does not find profitable to claim that he is a consumer. Since the value function Z is linear, these conditions simplify to

$$u\left(q^*\right) + \phi_b L \geq \phi_b B \tag{9}$$

$$-q_s^* + \phi_b K \geq \phi_b B \tag{10}$$

$$B \geq L \tag{11}$$

In addition, *participation constraints* require that producers, consumers, and inactive CCI members respectively, are better off using credit than cash; i.e.,

$$\begin{aligned} -q_s^* + Z(K,m) &\geq -q_{s,m} + \max\left\{Z(B, m + pq_{s,m}), V(m + pq_{s,m})\right\}, \\ u(q^*) + Z(L,m) &\geq u\left(q_m\right) + \max\left\{Z\left(B, m - pq_m\right), V\left(m - pq_m\right)\right\}, \\ Z(B,m) &\geq \max\{Z(B,m), V(m)\}. \end{aligned}$$

where m can be zero. The first constraint states that a seller is better off producing the efficient amount and incurring balance adjustment K than reporting inactivity, incurring adjustment B (if he stays in the credit arrangement) and selling his good for cash instead. The second constraint states that a consumer is better off consuming the efficient quantity using credit than reporting inactivity (or opting out of the credit arrangement) and using cash. The third condition, the participation constraint for the inactive agents, is just a special case of the no-default condition (7) and drops out. Note that condition (7) implies that if they use cash, active credit agents will prefer to stay in the credit arrangement. Exploiting the linearity of the value function Z, the nonredundant participation constraints can be rewritten as

$$\begin{aligned} -q_s^* + \phi_b (K - B) &\geq -q_{s,m} + \phi p q_{s,m} \\ u(q^*) + \phi_b (L - B) &\geq u(q_m) - \phi p q_m. \end{aligned}$$

Simplifying these expressions using $p\phi = 1$, these reduce to

$$-q_s^* + \phi_b (K - B) \geq 0 \tag{12}$$

$$u(q^*) + \phi_b (L - B) \geq u(q_m) - q_m. \tag{13}$$

Now, if (12) and (13) hold then (10) and (9) hold as well.

As before we confine our attention to stationary equilibria and require that $\phi_{+1} M_{+1} = \phi M$. We also require that $\phi_b X = \phi_{b,+1} X_{+1}$, where X denotes any balance adjustments. In the following we will normalize $\phi_b = 1$ and consider constant balance adjustments. We are now in a position to define a cash-credit equilibrium and state the main results of the paper (proofs are in the Appendix).

Definition 2 *A stationary cash-credit equilibrium is a credit system* (L, K, B, q^*) *and a list* (γ, q_m) *satisfying the (1),(4), (7), (8), (11), (12) and (13).*

In words, a cash-credit equilibrium must satisfy the conditions for a stationary monetary equilibrium, as well as the no-default, cost-recovery, incentive, and participation constraints necessary to sustain the CCI. In such an equilibrium, only the zero-measure group of infinite-lived agents transacts with cash. All other transactions occur through the CCI.

Proposition 2 *A stationary cash-credit equilibrium exists if* $\mu < \overline{\mu}$ *and* $\delta \leq \overline{\gamma}(\mu) \leq \gamma$, *where* $\overline{\gamma}'(\mu) > 0$.

Therefore, there is an equilibrium where credit coexists with cash as long as the monitoring cost is low enough. Indeed, if the cost of monitoring agents were too large (say infinite), then the financing of the credit arrangement would violate the participation constraint in the credit arrangement, independently of the value of cash. However, when the participation constraint is satisfied for low monitoring costs, the value of cash then matters by affecting the no-default constraint (7). The intuition is as follows. As the monitoring cost increases, participants in the credit arrangement have to contribute more to the credit arrangement in each period. Their incentive to use cash then increases. The credit arrangement will then only exist for relatively high values of γ, i.e., high implicit costs of holding cash. As γ increases, then from (4) the value of cash decreases and the level of consumption obtainable with cash decreases, making the credit arrangement more attractive.

Proposition 3 *Credit increases welfare if* $\mu < \mu^* \equiv a\,(1-n)\,[u\,(q^*) - q^*] - \delta q^*$.

The proof for this proposition considers the case where $\gamma = \delta$, so that cash is costless to hold for all agents. In this case only young agents are penalized in a cash economy, as they cannot consume in their first trading period. However, if the monitoring cost is low enough, young agents are better off ex-ante if they access the credit arrangement and then amortize the cost of the monitoring over subsequent periods, than by foregoing the chance to consume in their first period of life. The threshold μ^* is given by the the expected gain from participating in the trading stage in the first period of life, minus the cost of acquiring cash in the next settlement stage to consume q^*, that is, $\gamma q^* = \delta q^*$.

The threshold value μ^* obviously depends on the probability of consuming in the trading stage: if this probability is too low, then μ^* is negative and credit does not increase welfare.

4.3 The No-Surcharge Rule

A version of NSR arises quite naturally from the cash-credit equilibrium described above. We say that a cash-credit equilibrium follows a *no-surcharge rule* when a consumer's per-unit cost of purchasing a specialized good through the CCS does not exceed his cost of making the same purchase using cash.[12] Expressing consumers' cost of a credit purchase of q^* specialized goods (and so incurring balance L) in terms of the numeraire good, then NSR holds if

$$-\frac{L}{q^*} \leq p\phi$$

Recalling that in equilibrium $p\phi = 1$ and rearranging, this reduces to

$$-q^* \leq L \tag{14}$$

In the Appendix we that that if the Friedman rule is in effect so that $\gamma = \delta$ and $q_m = q^*$, then the no-default condition (7) reduces to

$$\delta(\mu - q^*) \leq L \tag{15}$$

which implies the no-surcharge rule (14). Since $-q^* < \delta(\mu - q^*)$, then by continuity, no-surcharge must also hold for rates of money growth slightly larger than δ. We state this as

Corollary 4 *In the cash-credit equilibrium, the no-surcharge rule (14) must hold for γ sufficiently close to δ.*

At higher money growth rates, however, no-surcharge may fail. In other words, no-surcharge is needed exactly when the implicit cost of using money is low, and some enticement is necessary to induce agents to transact through the CCS.

The cash-credit equilibrium may also require that producers of specialized goods, in effect, pay a form of merchant fee when they receive credit payments, i.e., that they receive less compensation per unit sold than producers who sell for cash. Analogous

[12] By stating this rule as an inequality we allow for the possibility of paybacks for card use.

to no-surcharge condition (14), we can say that a merchant fee is charged when a producer obtains less by selling on credit (and so obtaining balance K) than he would have by selling for cash, i.e., when

$$\frac{nK}{1-n} < 1$$

To see that a merchant fee can be charged in equilibrium, note that cost-recovery condition (8), combined with incentive constraint (11), places an upper bound on the compensation K of producers who sell for credit, as measured in terms of the numeraire

$$K \leq -\frac{1-\delta}{an}\mu - \frac{1-an}{an}L \tag{16}$$

When the Friedman rule is in effect, we can then use (15) to get the following upper bound on the credit producer's compensation

$$K \leq \frac{1-an}{an}\delta q^* - \frac{1-\delta}{an}\mu \tag{17}$$

Producers who sell for cash obtain

$$p\phi\frac{1-n}{n}q^* = \frac{1-n}{n}q* \tag{18}$$

Thus, merchants pay a merchant fee to receive credit payments if RHS(17) $<$ RHS(18), which can clearly occur for a close to 1.

These calculations demonstrate that a cash-credit equilibrium in the model can mimic the seemingly paradoxical real-world preference for card payments over cash. For the cases considered, adult consumers who have a low-cost alternative to the credit arrangement (i.e., cash) still have an incentive to pay by credit, since the price of paying by credit is no more than paying by cash. Producers agree to sell goods on credit, even though they receive less (in numeraire terms) than they would if they sold for cash.

The key to this somewhat magical arrangement is the possibility that an agent may be in an inactive state during the trading stage, combined with the agent's private knowledge of his state. The possibility of inactivity is meaningful because it implies that nonparticipation in the trading-stage market does not necessarily coincide with defection from the credit arrangement. Agents can be induced to truthfully reveal themselves as consumers, however, by in effect "charging a fee" to inactive agents, thereby keeping consumers' price of credit purchases low. This in turn discourages consumers from defaulting and going over to cash. Inactive agents and producers in our model continue to participate because they realize that at some point they will benefit as consumers.

The critical role of the inactive state can be illustrated if we now suppose the Friedman rule holds and simultaneously drive $a \to 1$, so that agents are always active during the trading stage. In the case we must set the inactive agents' balance $B = 0$ in conditions (9)-(13). Under the Friedman rule, the producers' participation constraint (12) reduces to

$$L \leq -q^* - \frac{1-\delta}{1-n}\mu \tag{19}$$

which is inconsistent with the no-default condition (15). Not coincidentally, (19) also violates the no-surcharge condition (14); adult agents would have no incentive to keep making credit purchases, and the credit arrangement collapses.

To summarize, in our model NSR plays the socially valuable role of making people demand credit who don't need it. The use of credit by adults who are *not* credit-constrained is beneficial to the young agents who are. No-surcharge is a way of bribing the adults into supporting the young. No-surcharge would not matter without limited enforcement: if adults could be compelled to transact through the CCI, the no-default constraint (7) would drop out and the planner would have more flexibility about how to allocate the costs of the credit arrangement.

5 Literature review

The approach outlined above follows the papers in the money literature that model credit arrangements as clubs, where membership in a club implies mutual knowledge of club members' identities and histories (or a sufficient subset thereof).[13] As in many of these papers, our model also allows club members the option of transacting anonymously with cash, which serves to tighten members' participation constraints. In such models, if money is divisible then credit arrangements become difficult or impossible to sustain when the monetary policy follows the Friedman rule. Our model avoids this fate by introducing young (credit-constrained) agents, and by restricting the ability of the central bank to make transfers to these agents. Thus, as in Freeman's (1996) overlapping-generations model of payments, credit is still useful even when cash is cheap.

Paralleling other papers in this literature, "network effects" arise quite naturally in our model, since membership in (and repeated use of) the credit arrangement amounts to a kind of club good. What is new is that we show how a no-surcharge rule can be instrumental in supporting the credit arrangement, by causing agents to internalize the gains of club participation. This result will hardly surprise people familiar with the literature on the industrial organization of the payment card industry, where network effects have been a dominant theme of discussion (e.g., Rochet and Tirole 2006). As noted in the introduction, however, our emphasis here is not on industrial organization, but on understanding the role of NSR in the context of monetary theory.

Telyukova and Wright (2006) employ a model similar to ours to explain another puzzling aspect of credit cards, which is why people do not use their non-interest-bearing cash to pay off interest-bearing credit-card balances. They show that this situation can persist if buyers are credit-constrained (so that credit cards have value), but cash is more widely accepted than cards (so cash has value). Again our focus is somewhat different, as we try to provide an explanation of how credit cards might displace cash under an optimal payments arrangement, even for purchases by buyers who are not credit constrained.

[13] Including Aiyagari and Williamson (2000), Corbae and Ritter (2004), Kahn and Roberds (2005), and Martin, Orlando, and Skeie (2006). In Berentsen, Camera, and Waller (forthcoming) this information is managed by specialized intermediaries (banks).

6 Conclusion

Above we have presented an environment where some type of payment system is needed for exchange. Fiat money allows some trades to occur but does not attain a first best allocation, even when the central bank follows the best possible policy, the Friedman rule. A credit-based payment system can improve on allocations attainable by trading only with cash. For monetary policies that are close to the Friedman rule, a no-surcharge rule may be necessary to ensure the viability of the credit-based system.

More generally we have attempted to illustrate how the tools of monetary theory can be applied to the analysis of payment systems. Although our model is quite stylized, it also highlights the key services provided by these systems–identification (authentication) and recordkeeping–in a fully dynamic, general-equilibrium environment. This approach may prove a useful one in exploring the nature of the benefits such systems provide, as well as the many policy issues associated with their operation.

7 Appendix: proofs

7.1 Derivation of the no-default condition

By definition, agents with balance $b \in \{K, L, B\}$ do not default whenever

$$Z(b, m) \geq \phi(m - m_{+1}) + \delta W(m_{+1}).$$

The inactive agents' incentive constraint (11) imposes $B \geq L$. Also, from the incentive constraint for sellers (10), we have $K > B$. Hence, consumers receive the minimum balance adjustment from participation in the CCS. It is therefore enough to consider the no-default condition for consumers, or

$$Z(L, m) \geq \phi(m - m_{+1}) + \delta W(m_{+1}). \tag{20}$$

From the linearity of Z in m we have

$$\phi m + Z(L, 0) \geq \phi(m - m_{+1}) + \delta W(m_{+1}),$$

so that the no-default constraint (20) becomes

$$Z(L, 0) \geq -\phi m_{+1} + \delta W(m_{+1})$$

Again using linearity of Z, this is

$$\phi_b L + \delta H(0, 0) \geq -\phi m_{+1} + \delta W(m_{+1})$$

From the budget constraint of cash buyers in the trading stage we have $m_{+1} = p_{+1} q_m$. Also, from (3) we have $\phi_{+1} p_{+1} = 1$. Therefore $\phi m_{+1} = \frac{\phi}{\phi_{+1}} \phi_{+1} p_{+1} q_m = \gamma q_m$. So condition (20) becomes

$$\phi_b L + \delta H(0, 0) \geq -\gamma q_m + \delta W(m_{+1})$$

Using linearity of H this is

$$\phi_b L + \frac{\delta}{1-\delta}\{a(1-n)[u(q^*) - q^*] + a\phi_b[nK + (1-n)L] + (1-a)\phi_b B\} \geq$$

$$-\gamma q_m + \delta W(m_{+1}) \Leftrightarrow$$

$$\phi_b L + \frac{\delta}{1-\delta}\{a(1-n)[u(q^*) - q^*] - (1-\delta)\mu\} \geq -\gamma q_m + \delta W(m_{+1})$$

where the last inequality follows from (8). Therefore we may restate (20) as

$$\begin{aligned}
\phi_b L - \delta\mu + \frac{\delta a (1-n)}{1-\delta}\left[u(q^*) - q^*\right] &\geq -\gamma q_m + \delta W(m_{+1}) \\
&\geq -\gamma q_m + \frac{\delta a (1-n)}{1-\delta}\left[u(q_m) - q_m\right]
\end{aligned}$$

To derive the second inequality, we used the fact that the expected hours worked on the settlement stage for a cash agent are by market clearing $a\left[(1-n)h_b + nh_s\right] + (1-a)h_i = 0$. Rearranging terms, the no-default condition is then given by

$$\phi_b L - \delta\mu \geq -\gamma q_m + \frac{\delta a (1-n)}{1-\delta}\left[u(q_m) - u(q^*) - (q_m - q^*)\right]$$

7.2 Proof of Proposition 2

Proof. Normalizing $\phi_b = 1$ we have two equilibrium equations characterizing the cash side of the economy:

$$\begin{aligned}
\frac{\gamma - \delta}{\delta} &= (1-n)\left[u'(q_m) - 1\right] \\
\gamma &= 1 + \tau.
\end{aligned}$$

Hence, we need $\gamma \geq \delta$. The constraints on the credit side of the economy are

$$\begin{aligned}
B &\geq L \\
-\frac{(1-n)}{n}q^* + K - B &\geq 0 \\
u(q^*) + L - B &\geq u(q_m) - q_m \\
L - \delta\mu &\geq -\gamma q_m + \frac{\delta a (1-n)}{1-\delta}\left[u(q_m) - u(q^*) - (q_m - q^*)\right] \\
a\left[nK + (1-n)L\right] + (1-a)B &= -(1-\delta)\mu
\end{aligned}$$

From these constraints it follows that setting $B = L$ slackens all the other constraints, thus increasing the set of parameters for which a cash-credit equilibrium exists. Hence we set $B = L$ in what follows. The inequalities for the credit side of the economy then become

$$-\frac{(1-n)}{n}q^* + K \geq L \tag{21}$$

$$u(q^*) \geq u(q_m) - q_m \tag{22}$$

$$L - \delta\mu \geq -\gamma q_m + \frac{\delta a(1-n)}{1-\delta}\left[u(q_m) - u(q^*) - (q_m - q^*)\right] \tag{23}$$

$$a\left[nK + (1-n)L\right] + (1-a)L = -(1-\delta)\mu \tag{24}$$

Condition (22) is always satisfied and is therefore redundant. From the cost-recovery condition we get an expression for K as a function of L.

$$K = -\frac{(1-an)L + (1-\delta)\mu}{an} \tag{25}$$

(Note that the merchant fee measured in terms of the numeraire is $\phi p - K/\left[\frac{(1-n)}{n}q^*\right] = 1 - nK/\left[(1-n)q^*\right]$.) Substituting (25) in (21) we can reduce the constraints for the credit arrangement to

$$\begin{aligned} -\frac{(1-n)}{n}q^* - \frac{(1-an)L + (1-\delta)\mu}{an} &\geq L \\ L - \delta\mu &\geq -\gamma q_m + \frac{\delta a(1-n)}{1-\delta}\left[u(q_m) - u(q^*) - (q_m - q^*)\right] \end{aligned}$$

Arranging terms we get

$$\begin{aligned} -a(1-n)q^* - (1-\delta)\mu &\geq L \\ L - \delta\mu &\geq -\gamma q_m + \frac{\delta a(1-n)}{1-\delta}\left[u(q_m) - u(q^*) - (q_m - q^*)\right] \end{aligned}$$

Therefore an equilibrium exists if

$$-a(1-n)q^* - (1-\delta)\mu \geq L \geq \delta\mu - \gamma q_m + \frac{\delta a(1-n)}{1-\delta}\left[u(q_m) - u(q^*) - (q_m - q^*)\right]$$

or if

$$-a(1-n)q^* - \mu \geq -\gamma q_m + \frac{\delta a(1-n)}{1-\delta}\left[u(q_m) - u(q^*) - (q_m - q^*)\right]$$

which is equivalent to

$$a(1-n)q^* + \mu - \frac{\delta a(1-n)}{1-\delta}\left[u(q^*) - q^*\right] \leq \gamma q_m - \frac{\delta a(1-n)}{1-\delta}\left[u(q_m) - q_m\right]. \tag{26}$$

When $\gamma > \delta$, the left hand side of (26) is constant in γ, while the derivative of the right hand side is

$$q_m + \gamma \frac{dq_m}{d\gamma} - \frac{\delta a (1-n)}{1-\delta} \left[u'(q_m) - 1 \right] \frac{dq_m}{d\gamma}$$

where using the equilibrium condition on the money market

$$\frac{\gamma - \delta}{\delta} = a(1-n)\left[u'(q_m) - 1\right]$$

we have

$$\begin{aligned} 1 &= \delta a (1-n) u''(q_m) \frac{dq_m}{d\gamma}, \text{ and} \\ \frac{dq_m}{d\gamma} &= \frac{1}{\delta a (1-n) u''(q_m)} < 0 \end{aligned}$$

Therefore, the derivative of the right hand side is

$$\begin{aligned} & q_m + \gamma \frac{dq_m}{d\gamma} - \frac{\delta}{1-\delta} \frac{\gamma - \delta}{\delta} \frac{dq_m}{d\gamma} \\ = \; & q_m + \left(\gamma - \frac{\gamma - \delta}{1-\delta} \right) \frac{dq_m}{d\gamma} \\ = \; & q_m + \delta \left(\frac{1-\gamma}{1-\delta} \right) \frac{dq_m}{d\gamma} \\ = \; & q_m + \delta \left(\frac{1-\gamma}{1-\delta} \right) \frac{1}{\delta a (1-n) u''(q_m)} \\ = \; & q_m + \left(\frac{1-\gamma}{1-\delta} \right) \frac{1}{a (1-n) u''(q_m)} \end{aligned}$$

Now

$$\gamma = \delta a (1-n) \left[u'(q_m) - 1\right] + \delta$$

so the derivative of the right hand side is

$$\begin{aligned} q_m + \left(1 - \frac{\delta a (1-n) \left[u'(q_m) - 1\right] + \delta}{1-\delta} \right) \frac{1}{a (1-n) u''(q_m)} & = \\ q_m - \frac{\delta \left[u'(q_m) - 1\right]}{(1-\delta) u''(q_m)} + \frac{1 - 2\delta}{a (1-n) u''(q_m)} & > 0 \end{aligned}$$

which is guaranteed if δ is higher than $1/2$. Therefore under this condition the RHS of (26) is increasing in γ. Now, at $\gamma = \delta$, we have $q_m = q^*$ so that (26) becomes

$$a(1-n) q^* + \mu \leq \delta q^*$$

This condition is satisfied if δ is close enough to one and μ is small enough. If this condition holds for some μ, then (26) holds for all $\gamma > \delta$. However if the above condition does not hold, then γ must be increased. Then either there is a $\bar{\gamma}(\mu)$ high enough so that (26) holds with equality (and therefore holds with strict inequality for all $\gamma > \bar{\gamma}(\mu)$), or (26) never holds (this is the case if for instance μ is large and the RHS of (26) reaches an asymptote as $\gamma \to \infty$, which depends on the third derivative of the utility function). Therefore, there is a μ' such that for all $\mu < \mu'$, (26) holds if and only if $\gamma \geq \overline{\gamma}(\mu)$. Furthermore $\overline{\gamma}'(\mu) > 0$.

Finally, a cash-credit equilibrium will not exist if the expected payoff from participating in the credit scheme is negative, that is if

$$a(1-n)[u(q^*) - q^*] - (1-\delta)\mu \leq 0,$$

or if

$$\widehat{\mu} = \frac{a(1-n)}{(1-\delta)}[u(q^*) - q^*] \leq \mu.$$

Therefore a cash-credit equilibrium exists if and only if $\mu \leq \overline{\mu} = \max\{\widehat{\mu}, \mu'\}$. ■

7.3 Proof of Proposition 3

Proof. Credit increases welfare relative to cash whenever the expected welfare in the credit economy is higher than the expected welfare in the cash economy, i.e., when

$$\frac{1}{1-\delta}\{a(1-n)[u(q^*) - q^*] - (1-\delta)\mu\} \geq an(-q_s^y + \phi p q_s^y) + \phi m_{+1} + \frac{\delta}{1-\delta}a(1-n)[u(q^*) - q^*]$$

Recalling that $p\phi = 1$ and $\phi m_{+1} = \gamma q_m$, this is equivalent to

$$\frac{1}{1-\delta}\{a(1-n)[u(q^*) - q^*] - (1-\delta)\mu\} \geq \gamma q_m + \frac{\delta}{1-\delta}a(1-n)[u(q^*) - q^*]$$

Which, under the best-case scenario for cash (Friedman rule) is the same as

$$\begin{aligned} \frac{1}{1-\delta}\{a(1-n)[u(q^*) - q^*] - (1-\delta)\mu\} &\geq \delta q^* + \frac{\delta}{1-\delta}a(1-n)[u(q^*) - q^*] \Leftrightarrow \\ (1-\delta)a(1-n)[u(q^*) - q^*] - (1-\delta)\mu &\geq (1-\delta)\delta q^* \end{aligned}$$

Therefore credit dominates cash for

$$\mu > \mu^* \equiv a(1-n)[u(q^*) - q^*] - \delta q^*$$

■

7.4 Derivations of NSR and merchant fees

Recall the NSR holds if

$$-q^* \leq L$$

Using the lower bound on L from (26), this will hold when

$$\delta\mu - \gamma q_m + \frac{\delta a(1-n)}{1-\delta}[u(q_m) - u(q^*) - (q_m - q^*)] \geq -q^*$$

setting $\gamma = \delta$, this reduces to

$$\begin{aligned} \delta\mu - \delta q^* &\geq -q^* \Leftrightarrow \\ (1-\delta) q^* &\geq -\delta\mu \end{aligned}$$

which is always true.

Let us now check the upper bound value for K. Again using the lower bound on L we have

$$\begin{aligned} anK &= -(1-\delta)\mu - (1-an) L \\ &\leq -(1-\delta)\mu - (1-an)\left[\delta\mu - \gamma q_m + \frac{\delta a(1-n)}{1-\delta}[u(q_m) - u(q^*) - (q_m - q^*)]\right] \Leftrightarrow \\ anK &\leq -(1-an\delta)\mu - (1-an)\left[-\gamma q_m + \frac{\delta a(1-n)}{1-\delta}[u(q_m) - u(q^*) - (q_m - q^*)]\right] \end{aligned}$$

When $\gamma = \delta$, the last inequality becomes

$$K \leq \frac{(1-an)}{an}\delta q^* - \frac{(1-an\delta)}{an}\mu$$

When they sell for cash, producers get $p\frac{(1-n)}{n}q^*$ which equals in terms of the numeraire good $\phi p\frac{(1-n)}{n}q^* = \frac{(1-n)}{n}q^*$. But since

$$\begin{aligned} \frac{(1-an)}{an}\delta q^* - \frac{(1-an\delta)}{an}\mu &\leq \frac{(1-n)}{n}q^* \Leftrightarrow \\ (1-an)\delta q^* - (1-an\delta)\mu &\leq a(1-n) q^* \end{aligned}$$

holds for a close enough to 1 and $\mu > 0$, producers paid by credit are receive less than those who are paid by cash (note however that a cash producer must still make a payment to the CCS as an "inactive" agent to remain within the CCS). Under the no-default condition, however, producers are still willing to participate in the CCS, and receive payment through the CCS, as they will benefit from the CCS when they are consumers.

References

[1] Aiyagari, S. Rao and Stephen D. Williamson, 2000. "Money and Dynamic Credit Arrangements with Private Information," *Journal of Economic Theory*, 91 (2), 248-79.

[2] Berentsen, Aleksander, Gabriele Camera, and Christopher Waller, (forthcoming). "Money, Credit and Banking," forthcoming in the *Journal of Economic Theory.*

[3] Chakravorti, Sujit, 2003. Theory of Credit Card Networks: a Survey of the Literature, *Review of Network Economics* 2, 50-68.

[4] Chakravorti, Sujit and Ted To, 1999. "A Theory of Credit Cards." Federal Reserve Bank of Chicago Working Paper 1999-16.

[5] Committee on Payment and Settlement Systems, 2006. *Statistics on payment and settlement systems in selected countries-Figures for 2004* (Basel, Bank for International Settlements).

[6] Corbae, Dean and Joseph Ritter, 2004. "Decentralized Credit and Monetary Exchange without Public Record Keeping," *Economic Theory*, 24, 933-951.

[7] Evans, David S., and Richard Schmalensee, 2005. "The Economics of Interchange Fees and Their Regulation: an Overview." Paper delivered at Interchange Fees in Credit and Debit Card Industries: What role for Public Authorities?, Federal Reserve Bank of Kansas City, May.

[8] Freeman, Scott, 1996. "The Payments System, Liquidity, and Rediscounting," *American Economic Review* 86, 1126-1138.

[9] Garcia-Swartz, Daniel D., Robert W. Hahn, and Anne Layne-Farrar, 2006. "The Move Toward a Cashless Society: A Closer Look at Payment Instrument Economics," *Review of Network Economics* 5(2), 175-198.

[10] Gerdes, Geoffrey, Jack K. Walton II, May X. Liu, and Darrel W. Parke, 2005. "Trends in the Use of Payment Instruments in the United States," *Federal Reserve Bulletin* 91, 180-201.

[11] Humphrey, David, Magnus Willesson, Ted Lindblom, and Göran Bergendahl, 2003. "What Does It Cost to Make a Payment?" *Review of Network Economics* 2(2), 159-174.

[12] Hunt, Robert M., 2003. "An Introduction to the Economics of Payment Card Networks," *Review of Network Economics* 2, 80-96.

[13] Kahn, Charles M. and William Roberds, 2005. "Credit and Identity Theft," Federal Reserve Bank of Atlanta Working Paper 2005-19.

[14] Klee, Elizabeth, 2006. "Families' Use of Payment Instruments during a Decade of Change in the U.S. Payment System," Board of Governors of the Federal Reserve System Finance and Economics Discussion Paper 2006-01.

[15] Kocherlakota, Narayana R., 1998. "Money is Memory," J*ournal of Economic Theory* 81, 232-251.

[16] Koeppl, Thorsten, Cyril Monnet, and Ted Temzelides, 2005. "A Dynamic Model of Settlement."

[17] Lagos, Ricardo, and Randall Wright, 2005. "A Unified Framework for Monetary Theory and Policy Analysis," *Journal of Political Economy* 113(3), 463-484.

[18] Lowe, Philip 2005. "Payments System Reform: the Australian Experience," Remarks prepared for the Federal Reserve Bank of Kansas City Conference on "Interchange Fees in Credit and Debit Card Industries: What role for Public Authorities?" Accessed online at www.kansascityfed.org/publicat/psr/proceedings/2005/lowe.pdf.

[19] Macfarlane, Ian J., 2005. "Gresham's Law of Payments," Speech at the AIBF Industry Forum. Accessed online at www.rba.gov.au/speeches/2005/sp_gov_230305.html.

[20] Martin, Antoine, Michael Orlando, and David Skeie, 2006. "Payment Networks in a Search Model of Money."

[21] Rochet, Jean-Charles and Jean Tirole, 2002. "Some Economics of Payment Card Associations," *Rand Journal of Economics* 33(4), 549-570.

[22] Rochet, Jean-Charles and Jean Tirole, 2004. "Two-sided Markets: an Overview." IDEI working paper, University of Toulouse.

[23] Rochet, Jean-Charles and Jean Tirole, 2006. "Externalities and Regulation in Card Payment Systems," *Review of Network Economics* 5(1): 1-14.

[24] Telyukova, Irina A., and Randall Wright, 2006. "A Model of Money and Credit, with Application to the Credit Card Puzzle."