WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 07-26

OPTIMAL PRICING OF PAYMENT SERVICES WHEN CASH IS AN ALTERNATIVE

Cyril Monnet
Federal Reserve Bank of Philadelphia

William Roberds
Federal Reserve Bank of Atlanta

October 18, 2007

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/econ/index.html

Optimal Pricing of Payment Services When Cash Is an Alternative *

Cyril Monnet
FRB Philadelphia
cyril.monnet@phil.frb.org

William Roberds
FRB Atlanta
william.roberds@atl.frb.org

October 18, 2007

Abstract

Payments are increasingly being made with payment cards rather than currency–this despite the fact that the operational cost of clearing a card payment usually exceeds the cost of transferring cash. In this paper we examine this puzzle through the lens of monetary theory. We consider the design of an optimal card-based payment system when cash is available as an alternative means of payment, and derive conditions under which cards will be preferred to cash. We find that a feature akin to the controversial "no-surcharge rule" may be necessary to ensure the viability of the card payment system. This rule, which is part of the contract between a 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.

*We thank audiences at the University of Rome Tor-Vergata, The Board of Governors, the Federal Reserve Banks of New York and Philadelphia, as well as Bob Hunt, Nobu Kiyotaki, Antoine Martin, Jamie McAndrews, Steve Williamson, and Randy Wright for comments and suggestions. The views expressed in the paper are not necessarily those of the Federal Reserve Bank of Philadelphia, the Federal Reserve Bank of Atlanta, or the Federal Reserve System. This paper is available free of charge at www.philadelphiafed.org/econ/wps/index.html

1 Introduction

Transactions that were once conducted only with cash–purchases of 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. Evidently, money is being displaced by memory.[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–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

[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 significant 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).

pay in many situations.

The higher cost of card payments is not readily apparent to purchasers because it is borne by merchants in the form of a *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 percent of the purchase amount for credit cards.[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 passing through this fee to 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 (cards) over 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 so easily in the case of debit cards or credit cards that are paid off every month. "Paybacks" to card use, commonly given in the form of frequent flyer miles, cash-back, or other rewards, further increase buyers' incentives to use cards while reducing the "social cost" of card transactions. Even so, paybacks cannot explain why a payments instrument with an apparently even lower social cost, cash, is not preferred.

As a benchmark, 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

[3] Average merchant fees on credit card transactions in 2005 were 2.19 percent for MasterCard and Visa, 2.41 percent for American Express, and 1.76 percent 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 discussion of this idea. Lowe (2005) notes that surcharging for card use is still uncommon in Australia, despite the regulatory removal of NSR.

in a payments environment, due to an asymmetry in market power between merchants and consumers.[6] A no-surcharge rule, it is argued, may actually improve welfare by preventing monopolistic merchants from inefficiently shifting the costs of a card payment system to consumers.

Below we consider the issue of card pricing as an application of monetary theory. We specify a dynamic environment that incorporates 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 market participants, and, once these have been identified, 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 card-based 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, and can be transferred for free. We then show how a planner would structure an optimal card system when cash is available. Certain features of real-world card payment systems, such as merchant fees and the no-surcharge rule, emerge endogenously as an advantageous features of such a system.

The intuition behind the model's version of the NSR is particularly simple. In our environment, the use of a card payment system relaxes cash constraints on potential purchasers. A no-surcharge rule promotes broad participation in the card payment system by, in effect, taxing the use of cash. Participation in a card payment system has both a private benefit (relaxing constraints on purchasers) and social benefit (sustaining the system that allows for relaxation of these constraints), and a no-surcharge rule allows agents to fully internalize the effects of their participation.

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

2 The Environment

Modeling choices. The analysis obviously requires an environment where agents may pay either with cash or with a card. To provide some minimal degree of verisimilitude, payment cards in particular should incorporate both a "credit" or record-keeping function (Kocherlakota 1998) and a "payment" function–identification of otherwise anonymous transactors (Kahn and Roberds, forthcoming). That is, the cards in the model should, just as with actual payment cards, serve the dual purpose of recording agents' transaction histories and correlating transactors with histories.

Verisimilitude also requires that debt incurred through use of payment cards be subject to some limitations on enforcement. Card debt is rarely collateralized and by its nature somewhat risky. While most credit card transactions are eventually paid off (default rates average about 4 percent in the U.S.), very little of the defaulted debt is ever collected.[7] The need to incorporate limited enforcement sets our analysis off from standard cash-good/credit-good approaches in the macro literature, e.g., Lacker and Schreft (1996), where this friction does not arise.

The Lagos-Wright (2005) model is an attractive starting point for our study of the NSR, since it provides a tractable model of anonymous exchange where agents can hold arbitrary money balances. Berentsen, Camera, and Waller (2007) have shown how the Lagos-Wright environment can be modified to incorporate uncollateralized credit, and our analysis will follow their setup in many respects. Our approach is also quite similar to that of Telyukova and Wright (forthcoming), who study the credit card debt puzzle, i.e., why people simultaneously carry cash balances and credit card debt. We simplify and enrich their model slightly as we are interested in a different question, namely the pricing of card-based payment services.

Finally, to keep track of surcharges it is important that someone will always use cash. To this end our setup includes a group of agents who, by assumption, are excluded from making use of alternative payment arrangements.

Agents, preferences, and technology. Time $t = 0, 1, ...,$ is infinite and discrete. The economy consists of a $[0, 1]$ continuum of agents. Agents are infinitely lived and discount the

[7] 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).

future at rate β. A proportion δ of agents can be identified at a utility cost μ, in the sense that associated with each of these agents is a unique verifiable quantity that we will call his "identity."[8] For a proportion $(1-\delta)$ of the population, this verification cost is effectively infinite.

Agents produce and consume two types of nonstorable goods: a specialized good and a numeraire good.[9] Utility (disutility) of consumption (production) of the numeraire 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 subperiods, each with its own market. Markets in both subperiods are Walrasian in the sense that no agent has market power.[10] 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*.

In the second subperiod, 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 randomly selected proportion $a \in (0,1)$ of the population has the opportunity to actively engage in trade. Of these, a randomly selected proportion

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

[9] Models in the payment cards literature customarily divide agents ex-ante into producers (merchants) and consumers. Our approach here follows the tradition of the money literature in which each agent's role in decentralized markets is randomly determined and private information to the agent. We would still obtain the no-surcharge rule if agents were divided ex ante into producers and consumers.

[10] Our environment could be modified to allow for other pricing mechanisms at the transaction stage, such as price posting or bargaining (see Rocheteau and Wright 2005). Here we focus on Walrasian markets so as to abstract from issues considered in the industrial organization literature.

$1-n$ 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.

1. Settlement stage
Agents produce, trade, and consume numeraire good
2. Trading stage
2a. Agents learn if they are consumers, producers, or inactive
2b. Buyers (consumers) and sellers (producers) trade specialized goods
2c. Buyers consume specialized goods

Table 1: Events within a period

Allocations. In this economy, we will consider symmetric stationary feasible allocations. A stationary allocation is a vector (q, q_s, h_s, h_b, h_i). The triple (h_s, h_b, h_i) denotes hours worked for an agent of type j when he was a seller (s), buyer (b), or inactive (i) in the previous trading stage. An allocation is feasible if

$$(1-n)\,q - nq_s = 0, \tag{1}$$

$$a\,(nh_s + (1-n)\,h_b) + (1-a)\,h_i = 0. \tag{2}$$

where (1) is the market clearing condition for the specialized good in the trading stage and (2) 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 an agent at the start of a trading stage is

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

We take the planner's objective function to be the maximization of the agents' lifetime payoff $\mathcal{W}$:

$$(1-\beta)\,\mathcal{W} = W.$$

Using the market clearing conditions, this reduces to

$$(1-\beta)\,\mathcal{W} = a\left[u(q) - q\right].$$

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

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

3 The Cash Economy

As agents are not automatically identifiable 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.

Let M_t be the per capita supply of cash in period t. Cash grows at the exogenously given and time independent rate γ, so that $M_{t+1} = \gamma M_t$. In the analysis, we drop the index t when there is no chance of confusion, so that M_t becomes M and M_{t+1} becomes M_{+1}. We will concentrate the analysis on stationary monetary equilibria where $\phi_{+1}M_{+1} = \phi M$, ϕ being the real price of money in terms of numeraire. We will denote the growth rate of money M_{+1}/M by γ. Hence γ also equals ϕ/ϕ_{+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

$$\begin{aligned} V(m) &= \max_{h,m_{+1}} -h + \beta W(m_{+1}) \\ & \quad\; s.t. \quad \phi m_{+1} = h + \phi m + \phi\tau. \end{aligned}$$

where τ is a nominal transfer received by all agents. The first-order and envelope conditions give

$$\beta W'(m_{+1}) = \phi, \qquad V'(m) = \phi. \tag{3}$$

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.

Agents' discounted lifetime utility when they enter the trading stage with m units of cash is

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

where q and q_s are set optimally as follows. A *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 (3) we have

$$p\phi = 1. \tag{4}$$

A *consumer* solves

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

and the first-order condition gives

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

where $V'(m-pq)$ has been replaced by ϕ using (3). 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 $u'(q) = p\phi$.

It is now easy to determine the value of an additional unit of money when 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 $W'(m) = \phi_{-1}/\beta$. Hence we obtain another equilibrium condition, which characterizes the quantity consumed in the trading stage as a function of the money growth rate γ

$$\frac{\gamma-\beta}{\beta} = a(1-n)[u'(q)-1]. \tag{5}$$

We can now define and characterize a stationary monetary equilibrium.

Definition 1 *Given γ, a stationary monetary equilibrium is q such that (5) holds.*

Proposition 1 *A stationary monetary equilibrium exists for all $\gamma \geq \beta$, and is unique for all $\gamma > \beta$. The equilibrium consumption of the specialized good q is strictly decreasing in γ.*

The proof of the proposition follows immediately from condition (5). Note that if the Friedman rule holds, i.e., if $\gamma = \beta$, consumers always hold enough money to purchase the efficient amount, so that in equilibrium $u'(q) = 1$.

4 The Cash-Card Economy

The requirement to transact in cash constrains agents when monetary policy does not follow the Friedman rule. Away from the Friedman rule, the availability of other payment arrangements can potentially improve welfare by relaxing agents' cash constraints. In this section, we consider how one such arrangement affects welfare. This part of the paper draws on the private information approach developed in Koeppl, Monnet, and Temzelides (KMT, forthcoming and 2007).

Note that 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, a payment arrangement must also incorporate some technology for identifying debtors. To fix ideas, we imagine that this identification occurs using plastic cards.

More specifically, we may imagine that the planner relies on a club arrangement known as a *card club* (CC).[11] Each of the δ agents whose identity is verifiable may choose to participate in the club. The CC incurs monitoring costs $\mu \geq 0$ per club member, per period, in verifying the identity of its members.[12] This cost is denominated in the numeraire good

[11] A club arrangement is quite natural since verification of an agent's identity will amount to provision of a nonrival good. Reliance on the club implicitly assumes that the planner is solving a constrained Pareto problem of maximizing the welfare of those within the club.

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

and is incurred at the end of the settlement stage.[13] Once it has incurred this cost, the CC can verify agents as members and record a member's transactions during the trading stage. 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. Nor can the CC force continued participation in the club. Recall that the identity of the remainder $(1 - \delta)$ of the population is unverifiable, which prevents these agents from joining the CC.

The CC will seek to implement the first best allocation q^* in the trading stage for the members of the CC. Hence, CC does not seek to maximize profit but only to achieve q^* and to recover costs. Following KMT, we will design the terms of the card club so that agents truthfully reveal their state (consumer, producer, inactive) and so that they are willing to participate in the card arrangement given the outside option of using cash. The CC has no control over monetary policy and takes the money growth rate γ as given.

As in KMT, the CC assigns *balances* to participants.[14] The club 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.

Qualitatively, there are three possibilities regarding agents' trading-stage activity: 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

[13] At the cost of additional complexity, we could incorporate per-transaction costs of using the card system, as is customary in many models of card payments beginning with Baxter (1983). Here we abstract from these costs in order to focus on the key services provided by modern card payment systems–authentication and recordkeeping–whose marginal costs are close to zero under modern technology.

[14] These balances are precisely defined in KMT; for our purposes we will think of them as running totals of an agent's transactions.

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 CC cannot impose any direct penalty on a member with low balances who does not readjust his balance during a settlement period. That is, a CC member can walk away from the arrangement at any time. The only penalty the CC can apply is denial of future access.

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

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

Under assumption 1, KMT show that there exists a simple optimal CS, where balances upon exiting the settlement stage d_t are equal to zero for all t. In other words, if a member of the CC exits the settlement stage with $d \neq 0$, then the CC shuts down the account of this member so that his card becomes invalid. It follows that the equilibrium distribution of balances is degenerate at $d = 0$, when CC 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 CC of those eligible to use their cards. 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 CS. Then CC members decide if they will participate, and if so, whether they will use cash or cards. That is, they may participate anonymously as cash agents, or they may allow themselves to be identified, and transact using cards. CC members thus have the opportunity to use cash rather than cards at any time. Unverifiable agents remain outside the CC, and will always use cash. Hence, while an active cash market will always exist, a successful CS should contain incentives such that CC members transact using cards.

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 and those that use cards.[15] 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 might 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 may only transact with cash. We now describe the problem of agents with access to both cash and the CC.

4.1 Cards at the settlement stage

Let $Z(b,m)$ denote the value function of a CC 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}. If the price of balances in the settlement round is ϕ_b, CC members at the beginning of the settlement round solve the following:

$$\begin{aligned}Z(b,m) &= \max_{h,m_{+1}}\{-h+\beta H(b_{+1},m_{+1})\} \qquad (6)\\ \text{s.t.}\quad \phi m_{+1}+\phi_b b_{+1} &= h+\phi_b b+\phi m.\end{aligned}$$

The first order condition with respect to money gives

$$\beta H_m\left(0,m_{+1}\right)\leq\phi \text{ with strict inequality if } m_{+1}=0. \qquad (7)$$

The envelope conditions are $Z_b=\phi_b$ and $Z_m=\phi$. Since the CC requires card users to carry zero balances when they exit the settlement stage, under the threat of exclusion, we

[15] Some policy discussions of NSR have focused on the issue of potential cross-subsidization between purchasers using cash and those using cards. We abstract from this issue in order to focus on allocations across heterogeneous participants in the card arrangement.

have $b_{+1} = 0$. When CC 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 card balances and cash holdings.

Also, for any balance b, it should be the case that CC 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,

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

In the Appendix, we show that this requirement reduces to

$$\phi_b L - \mu \geq -\gamma q_m + \frac{\beta a (1-n)}{1-\beta} [u (q_m) - u (q^*) - (q_m - q^*)] \tag{8}$$

The *cost-recovery constraint* in the settlement stage requires that members cover the operational costs incurred by the CC

$$a\phi_b [nK + (1-n) L] + (1-a) \phi_b B + \mu = 0. \tag{9}$$

4.2 Cards at the trading stage

We now turn to the problem faced by CC 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 q_s^*. Consumers receive q^*. The auctioneer subsequently communicates the identity and reports of those CC members that made use of the CS to the CC, 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:

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

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

$$u(q^*) + \phi_b L \geq \phi_b B, \tag{10}$$

$$-q_s^* + \phi_b K \geq \phi_b B, \tag{11}$$

$$B \geq L. \tag{12}$$

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

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

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 card club) and selling his good for cash instead. The second constraint states that a consumer is better off consuming the efficient quantity using a card than reporting inactivity (or opting out of the card arrangement) and using cash. The third condition, the participation constraint for the inactive agents, is just a special case of the no-default condition (8) and drops out. Note that condition (8) implies that if they use cash, active card agents will prefer to stay in the club. 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

$$-q_s^* + \phi_b (K - B) \geq 0 \tag{13}$$

$$u(q^*) + \phi_b (L - B) \geq u(q_m) - q_m. \tag{14}$$

Now, if (13) and (14) hold then (11) and (10) 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-card equilibrium and state the main results of the paper (proofs are in the Appendix).

Definition 2 *Given γ, a stationary cash-card equilibrium is a card system (L, K, B, q^*) and a (cash) trading-stage consumption q satisfying (5), (8), (9), (12), (13), and (14).*

In words, a cash-card 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 CC. In such an equilibrium, only the $1 - \delta$ unverifiable agents transact with cash. All other transactions occur through the CC.

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

Therefore, there is an equilibrium where cards coexist with cash as long as the monitoring cost is low enough. Indeed, if the cost of monitoring agents were too large, then the financing of the card arrangement would violate the participation constraint in the card 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 (8). The intuition is as follows. As the monitoring cost increases, participants in the card arrangement have to contribute more to the card arrangement in each period. Their incentive to use cash then increases. The card arrangement will then only exist for relatively high values of γ, i.e., high implicit costs of holding cash. As γ increases, then from (5) the value of cash decreases and the level of consumption obtainable with cash decreases, making the card arrangement more attractive.

We next consider the issue of welfare.

Proposition 3 *Cards increase welfare if $\gamma \geq \tilde{\gamma}(\mu) > \beta$, where $\tilde{\gamma}'(\mu) > 0$.*

Not suprisingly, Proposition 3 shows that cards are welfare improving if cash is expensive, i.e., when the inflation tax is relatively high compared with the cost of using cards.

Some care is called for in applying this result, however. Under Proposition 2, a cash-card equilibrium can in some cases continue to exist (in steady state) even if the Friedman rule is implemented, and despite the risk of default.[16] Such an equilibrium would run counter to the rationale for the card club, however, since it would deliver the same allocation as cash but at a higher cost.

It is perhaps counterintuitive that a cash-card equilibrium can exist in cases where it results in lower welfare than the cash equilibrium. To see how this can happen, let us make the following thought experiment. Suppose a CC member defects from the club and reverts to using cash. Then this agent has to incur the cost of acquiring cash before returning to the cash economy. At the Friedman rule, it costs $\gamma q^* = \beta q^*$ to acquire the cash necessary to buy q^*. The discounted lifetime payoff for the defector is then $-\beta q^* + [u(q^*) - q^*] / (1-\beta)$. If this agent were to use a card instead, he would enjoy a lifetime payoff of $[u(q^*) - q^* - \mu] / (1-\beta)$. Hence, this agent should keep using his card, even at the Friedman rule, so long as $\mu < \beta(1-\beta) q^*$. In other words, switching to using cash is costly, as one needs to accumulate a big enough cash balance to consume. Therefore, although the steady-state payoff of using cash is higher than that of using cards, agents prefer to use cards as they do not need to accumulate the necessary cash balance.

A version of this same logic will be useful in our consideration of the no-surcharge rule, to which we now turn.

4.3 The no-surcharge rule

A version of NSR arises quite naturally from the cash-card equilibrium described above. We say that a cash-card equilibrium follows a *no-surcharge rule* when a consumer's per-unit cost of purchasing a specialized good through the card club does not exceed his cost of making the same purchase using cash.[17] Expressing consumers' cost of a card purchase of q^* specialized goods (and so incurring balance L) in terms of the numeraire good, then NSR holds if

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

[16] Precise conditions are given in the proof of Proposition 2 in the Appendix.

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

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

$$-q^* \leq L. \tag{15}$$

In the Appendix we show that if the Friedman rule is in effect so that $\gamma = \beta$ and $q_m = q^*$, and there is a cash-card equilibrium, then the no-default condition (8) takes on a simple form

$$\mu - \beta q^* \leq L. \tag{16}$$

In words, condition (16) states that the sum of the consumer's balance adjustment $-L$ and the monitoring cost μ cannot exceed the cost of acquiring the necessary cash to defect from the card arrangement, which is βq^* at the Friedman rule. Condition (16) clearly implies that the no-surcharge rule (15) must hold at the Friedman rule. And, since $-q^* < \mu - \beta q^*$, then by continuity, no-surcharge must also hold for rates of money growth slightly larger than β. We state this as

Corollary 4 *In the cash-card equilibrium, the no-surcharge rule (15) must hold for γ sufficiently close to β.*

Hence, for monetary policies sufficiently close to the Friedman rule, NSR is an integral feature of the card arrangement. At higher money growth rates, however, no-surcharge may not hold. 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 using cards.

The cash-card equilibrium may also require that producers of specialized goods, in effect, pay a form of merchant fee when they receive card payments, i.e., they receive less compensation per unit sold, in numeraire terms, than producers who sell for cash. Analogous to no-surcharge condition (15), we can say that a merchant fee is charged when a producer obtains less through a card sale (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 (9), combined with incentive constraint (12), places an upper bound on the compensation K of producers who make a card sale, as measured in terms of the numeraire

$$K \leq -\frac{1}{an}\mu - \frac{1-an}{an}L \tag{17}$$

When the Friedman rule is in effect, we can then use (16), i.e., $\beta q^* - \mu > -L$, to get the following upper bound on the card-accepting producer's compensation

$$\begin{aligned} K &\leq -\frac{1}{an}\mu - L\frac{1-an}{an} && (18) \\ &\leq -\frac{1}{an}\mu + (\beta q^* - \mu)\frac{1-an}{an} && (19) \\ &\leq \beta q^* \frac{1-an}{an} - \mu\frac{2-an}{an} && (20) \end{aligned}$$

Producers who sell for cash obtain

$$p\phi\frac{1-n}{n}q^* = \frac{1-n}{n}q* \tag{21}$$

Thus, producers pay a merchant fee to receive card payments if RHS(18) < RHS(21), which holds if

$$\beta q^* \frac{1-an}{an} - \mu\frac{2-an}{an} < \frac{1-n}{n}q*$$

which can clearly occur for a close to 1, or μ large enough. By continuity, the same conditions carry over to money growth rates sufficiently close to the Friedman rule.

These calculations demonstrate that a cash-card equilibrium in the model can mimic the seemingly paradoxical real-world preference for card payments over cash. For the cases considered, consumers who have a low-cost alternative to the card club (i.e., cash) still have an incentive to pay with a card, since the price of paying by card is no more than paying by cash. Producers agree to receive card payments, even though they receive less per unit sold (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 card 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 card 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 again consider the limiting case of the Friedman rule, and simultaneously set $a = 1$, so that agents are always active during the trading stage. In this case we must set the inactive agents' balance $B = 0$ in conditions (10)-(14). Under the Friedman rule, the producers' participation constraint (13) reduces to

$$L \leq -q^* - \frac{\mu}{1-n} \tag{22}$$

which is inconsistent with the no-default condition (16). Not coincidentally, (22) also violates the no-surcharge rule (15); agents would have no incentive to keep making card purchases, and the card club collapses.

To summarize, in our model NSR is sometimes necessary to ensure the viability of the club arrangement. At low inflation rates, without the no-surcharge rule the card equilibrium would collapse and only cash would be used. No-surcharge would not matter without limited enforcement: if agents could be compelled to transact through the CC, the no-default constraint (8) would drop out and the planner would have more flexibility about how to allocate the costs of card-based payment.

5 Discussion

The approach outlined above follows the papers in the money literature that model payment arrangements as clubs, where membership in a club implies mutual knowledge of club members' identities and histories (or a sufficient subset thereof).[18] 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 alternative payment arrangements become difficult to sustain as monetary policy approaches the Friedman rule. For the environment considered above, the no-default constraint (8) is crucial to the sustainability of the card club, and at low inflation rates this constraint approximates the no-surcharge rule.

When monetary policy follows the Friedman rule exactly, the model can sustain a sort of "Gresham's Law" equilibrium envisioned by Macfarlane (2005)–once a card arrangement is in place, agents will continue to use it even though, over the long run, they would be

[18] Including Aiyagari and Williamson (2000), Corbae and Ritter (2004), Kahn and Roberds (forthcoming), and Martin, Orlando, and Skeie (forthcoming). In Berentsen, Camera, and Waller (2007) this information is managed by specialized intermediaries (banks).

better off abandoning it in favor of cash. We do not regard this limiting result as evidence against the desirability of the no-surcharge rule. To the contrary, in the more applicable case where monetary policy does not follow the Friedman rule, our analysis shows that a policy of no surcharges on card purchases emerges as a natural feature of a welfare-enhancing arrangement. Thus, the mere existence of NSR in actual pricing schemes does not present prima facie evidence of suboptimal pricing.

Paralleling other papers in this literature, "network effects" arise quite naturally in our model, since membership in (and repeated use of) the card arrangement amounts to a kind of club good. We show how a no-surcharge rule can be instrumental in supporting the card arrangement, by allowing 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 card-based payments in the context of monetary theory.

6 Conclusion

Above we have presented an environment where some type of payment system is needed for exchange. Away from the Friedman rule, fiat money does not attain the first best allocation. A card-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 card-based system.

In the environment studied, no-surcharge is only valuable when there is both limited enforcement of debts and private information about cardholder's ability to supply a good, i.e., "repay." But since these frictions are pervasive in real-world payment situations, this restriction may be seen as more a desirable feature of the model than a limitation.

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 useful 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

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

The inactive agents' incentive constraint (12) imposes $B \geq L$. Also, from the incentive constraint for sellers (11), 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

$$Z(L, m) \geq \phi(m - m_{+1}) + \beta W(m_{+1}). \tag{23}$$

From the linearity of Z in m we have

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

so that the no-default constraint (23) becomes

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

Again using linearity of Z, this is

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

From the budget constraint of cash buyers in the trading stage we have $m_{+1} = p_{+1} q_m$. Also, from (4) 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 (23) becomes

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

Using linearity of H this is

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

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

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

where the last inequality follows from (9). Therefore we may restate (23) as

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

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

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

7.2 Proof of Proposition 2

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

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

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

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

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

$$-\frac{(1-n)}{n}q^* + K \geq L \tag{24}$$

$$u(q^*) \geq u(q_m) - q_m \tag{25}$$

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

$$a\left[nK + (1-n)L\right] + (1-a)L = -\mu \tag{27}$$

Condition (25) is always satisfied and is therefore redundant. From the cost-recovery condition we get an expression for K as a function of L.

$$K = -\frac{(1-an)\,L + \mu}{an} \tag{28}$$

(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 (28) in (24) we can reduce the constraints for the card arrangement to

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

Arranging terms we get

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

Therefore an equilibrium exists if

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

or if

$$-a\,(1-n)\,q^* - 2\mu \geq -\gamma q_m + \frac{\beta a\,(1-n)}{1-\beta}\left[u\,(q_m) - u\,(q^*) - (q_m - q^*)\right]$$

which is equivalent to

$$a\,(1-n)\,q^* + 2\mu - \frac{\beta a\,(1-n)}{1-\beta}\left[u\,(q^*) - q^*\right] \leq \gamma q_m - \frac{\beta a\,(1-n)}{1-\beta}\left[u\,(q_m) - q_m\right]. \tag{29}$$

When $\gamma > \beta$, the left-hand side of (29) is constant in γ, while the derivative of the right-hand side is

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

where using the equilibrium condition on the money market

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

we have

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

Therefore, the derivative of the right hand side is

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

Now

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

so the derivative of the right hand side is

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

which is guaranteed if β is higher than 1/2. Therefore under this condition the RHS of (29) is increasing in γ. Now, at $\gamma = \beta$, we have $q_m = q^*$ so that (29) becomes

$$a\,(1-n)\,q^* + 2\mu \leq \beta q^*$$

It is clear that this condition is satisfied for feasible parameters. For instance, if μ is sufficiently small and $a\,(1-n) < \beta$. If this condition holds for some μ, then (29) holds for all $\gamma > \beta$. 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 (29) holds with equality (and therefore holds with strict inequality for all $\gamma > \bar{\gamma}(\mu)$), or (29) never holds (this is the case if for instance μ is large and the RHS of (29) reaches an asymptote as $\gamma \to \infty$, which depends on the third derivative of the utility function). Therefore, there is a $\mu' > 0$ such that for all $\mu < \mu'$, (29) holds if and only if $\gamma \geq \overline{\gamma}\,(\mu)$. Furthermore $\overline{\gamma}'\,(\mu) > 0$.

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

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

or if

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

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

Remark 5 *In particular, a cash-card equilibrium will exist for* $\gamma = \beta$ *and*

$$\mu \leq \min\{a\,(1-n)\,[u\,(q^*) - q^*]\,, [\beta - a(1-n)]q^*/2\}$$

7.3 Proof of Proposition 3

Proof. Cards increase welfare relative to cash whenever the expected welfare in the card economy is higher than the expected welfare in the cash economy, i.e., when

$$a\,(1-n)\,[u\,(q^*) - q^*] - \mu \geq a\,(1-n)\,[u\,(q_m) - q_m]$$

It is clear that at the Friedman rule, cash dominates since $q_m = q^*$. However, $u\,(q_m) - q_m$ also tends to zero as γ tends to infinity. Hence, there is $\tilde{\gamma} > \beta$, such that the use of cards improves upon the use of cash. ■

7.4 Derivations of NSR and merchant fees

In this section we suppose that there exists a cash-card equilibrium even when $\gamma = \beta$. The extension to the other cases is straightforward, but cumbersome. Recall the NSR holds if

$$-q^* \leq L$$

Using the lower bound on L from (29), this will hold when

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

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

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

which is always true.

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

$$\begin{aligned} anK &= -\mu - (1-an)\,L \\ &\leq -\mu - (1-an)\left[\mu - \gamma q_m + \frac{\delta a\,(1-n)}{1-\delta}\left[u\left(q_m\right) - u\left(q^*\right) - \left(q_m - q^*\right)\right]\right] \Leftrightarrow \\ anK &\leq -\left(2-an\right)\mu - (1-an)\left[-\gamma q_m + \frac{\delta a\,(1-n)}{1-\delta}\left[u\left(q_m\right) - u\left(q^*\right) - \left(q_m - q^*\right)\right]\right] \end{aligned}$$

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

$$K \leq \frac{(1-an)}{an}\beta q^* - \frac{(2-an)}{an}\mu$$

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

$$\begin{aligned} \frac{(1-an)}{an}\beta q^* - \frac{(2-an)}{an}\mu &\leq \frac{(1-n)}{n}q^* \Leftrightarrow \\ (1-an)\,\beta q^* - (2-an)\,\mu &\leq a\,(1-n)\,q^* \end{aligned}$$

holds for a close enough to 1 and $\mu > 0$, producers paid by card are receive less than those who are paid by cash (note however that a cash producer must still make a payment to the CC as an "inactive" agent to remain within the CC). Under the no-default condition, however, producers are still willing to participate in the CC, and receive payment through the CC, as they will benefit from the CC 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] Baxter, W.F., 1983. "Bank Interchange of Transactional Paper: Legal Perspectives," *Journal of Law and Economics* 26, 541-588.

[3] Berentsen, Aleksander, Gabriele Camera, and Christopher Waller, 2007. "Money, Credit and Banking" *Journal of Economic Theory* 135(1), 171-195

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

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

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

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

[8] 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.

[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, forthcoming. "Credit and Identity Theft," forthcoming in the *Journal of Monetary Economics*.

[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, forthcoming. "A Dynamic Model of the Payment System." Federal Reserve Bank of Philadelphia, Working Paper 07-22.

[17] Koeppl, Thorsten, Cyril Monnet, and Ted Temzelides, forthcoming. "A Dynamic Model of Settlement." Forthcoming in the *Journal of Economic Theory.*

[18] Lacker, Jeffrey M. and Stacey L. Schreft, 1996. "Money and Credit as Means of Payment," *Journal of Monetary Economics* 38, 3-23.

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

[20] 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.

[21] 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.

[22] Martin, Antoine, Michael Orlando, and David Skeie, forthcoming. "Payment Networks in a Search Model of Money," forthcoming in the *Review of Economic Dynamics.*

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

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

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

[26] Rocheteau, Guillaume and Randy Wright, 2005. "Money in Search Equilibrium, in Competitive Equilibrium, and in Competitive Search Equilibrium," *Econometrica* 73: 175-202.

[27] Telyukova, Irina A., and Randall Wright, forthcoming. "A Model of Money and Credit, with Application to the Credit Card Puzzle," forthcoming in the *Review of Economic Studies.*