# From Fairs to Banks

Luis Araujo*
Michigan State University

Raoul Minetti
Michigan State University

**Abstract**

This paper puts forward a theory for the origin of banks based on their ability to record and disseminate information when trade is decentralized and information on transactions is scarce. We first show that when trade is limited cooperation in exchange can be sustained by institutions where transactions are publicly observable. We interpret these institutions as fairs. We then demonstrate that the expansion of trade can progressively undermine the ability of fairs to record and disseminate the information that is needed in exchanges. When trade intensifies, institutions can continue to support cooperation only if they evolve into banks and start issuing notes for the settlement of transactions. Crucially, the sustainability of bank-based exchange hinges on the same feature of institutions that initially supported fair-based exchange, that is their visibility, but it is not undermined by trade expansion. We argue that the model can help explain the decline of fairs and the emergence of banks in Medieval and Early Modern Europe. Moreover, it can offer insights into the challenges that the information revolution is posing to the role of banks in the payments system.

JEL Codes: D80, E00, G21.

*Keywords*: Banks, Notes, Information, Trade.

## 1 Introduction

The role of banks in facilitating the settlement of transactions is recognized as an engine of the secular development of trade.[1] In recent years, the information revolution has been increasingly challenging this role. The availability of electronic databases with detailed records of transactions, the growing use of internet in settlements, and the diffusion of e-money are jeopardizing the traditional importance of banks and bank instruments (e.g., cheques) in the payments system. In light

*Corresponding author. Address: Michigan State University, Economics Department, 110 Marshall-Adams Hall, East Lansing, MI, 48824-1038. For helpful comments and conversations, we are grateful to Braz Camargo, Leo Ferraris, Nobuhiro Kiyotaki, Lucy White, Randall Wright, and seminar and conference participants at the Federal Reserve Bank of Chicago Summer Workshop on Money, Banking and Payments, the Econometric Society Summer Meetings (Carnegie Mellon), the Midwest Economic Theory Meetings (University of Illinois at Urbana-Champaign), the Swiss Conference on Banking and Financial Intermediation (Champery), and Michigan State University. All remaining errors are ours.

[1] For an in-depth historical analysis see, e.g., De Roover (1948).

of this, exploring the origins of banks' role in the settlement of transactions - especially in the information flows involved in exchanges - is crucial to learn insights not only into the importance of banks but also into their future. Yet, the economic scholarship has produced relatively few formal studies on these origins. This paper puts forth an hypothesis based on the ability of banks to efficiently record and disseminate the information needed in trade. We demonstrate that, when trade expands and the amount of information necessary to sustain cooperation in the exchange process grows, institutions that sustain trade without the need of tangible media of exchange must evolve into banks, institutions that sustain trade by issuing such tangible instruments. Although a formalization of this conjecture is beyond the scope of this paper, we view the current transformations of the payments system as a reversal of this process. While trade is continuously expanding and the information required to support cooperation is growing, the advances in the technology for the recording and dissemination of information are eroding banks' importance in exchanges, prompting the re-emergence of institutions that sustain trade without the need of tangible instruments (e.g., electronic payment providers).[2]

Our hypothesis on the origins of banks can be summarized as follows. We study an economy where trade is decentralized and information on transactions is scarce. These features hinder cooperation in the exchange process: agents have the incentive to consume the goods produced by other agents without reciprocating with their own goods. In the economy, agents trade inside and outside institutions, the distinctive feature of institutions being that transactions occurring inside them are publicly observable. We show that when trade is limited institutions can sustain it because, whenever a non-cooperative behavior spreads through the economy, as soon as it reaches an institution it gets publicly disclosed. In this first stage, institutions thus resemble "fairs", locations where transactions occur and are recorded publicly. The ability of fairs to disseminate information about the lack of cooperation in exchanges hinges however on the physical occurrence and recording of transactions inside them. As trade expands, the ability of institutions to host and record transactions comes under strain. Therefore, institutions can continue to support trade only if they start acting as banks and issuing tangible objects ("notes") that guarantee the immediate payment of goods and, hence, contain all the information required in trade. The "visibility" of institutions continues to be key at this stage: it is because transactions inside them are publicly observable that institutions are discouraged from overissuing notes. Put differently, the same observability of transactions that initially allowed institutions to act as fairs (disseminating information upon agents' lack of cooperation) now allows them to act as banks (disseminating information upon attempts to overissue notes). We show that crucially, unlike in the case of fairs, the role of banks is not undermined by trade expansion because banks can always respond to this expansion by injecting more notes (liquidity) in the economy. In a sense, since it requires the physical occurrence and recording of transactions inside them, the technology for the dissemination of information that is used by fairs suffers from decreasing returns to scale when trade expands. In contrast, the technology that is used by banks does not

[2] Kocherlakota (1998) shows that fiat money is not essential as a medium of exchange if there is a technology that perfectly records the past actions of agents.

suffer from this problem. In fact, notes - which contain all the information required for transactions - are objects that can be issued at no cost; moreover, regardless of how intense trade is, only one piece of information - the total amount of notes - is needed to prevent banks' overissue of notes.

We argue that our hypothesis broadly matches historical evidence on the origin of banks in Medieval and Early Modern Europe. There is evidence that a sizable volume of public information and records was available about transactions occurring in fairs and banks during this period. This is consistent with our core assumption of institutions as locations where transactions are publicly observable. The progressive evolution from fairs to banks during this period is also well documented by historians, who argue that this evolution coincided with the expansion of trade. Indeed, the historical facts suggest that an important reason for the decline of fairs was their inability to keep up with the intensification of trade and the growing need of records and information associated with it.

The remainder of the paper is organized into five sections. In Section 2, we relate the paper to the prior literature. Section 3 presents and solves the model and explores robustness issues. Section 4 reviews historical facts that are consistent with our hypothesis. In Section 5, we discuss alternative interpretations of the tangible payment instrument used under bank-based trade and of the mechanism of exchange under fair-based trade. Section 6 concludes. Lengthy proofs are relegated to the Appendix.

## 2 Prior Literature

This paper contributes to the literature on the role and origin of banks. While scholars generally agree that banking arose from the evolution of mercantile activity, their explanations of this evolution differ from ours. Some scholars argue that the most successful merchants progressively accumulated capital in trades and at some point this enabled them to offer credit to other merchants. Other scholars focus on the accumulation of human capital rather than financial capital: by engaging in trade, merchants acquired skills and knowledge that eventually allowed them to perform the sophisticated activities of bankers (for a detailed discussion of these two hypotheses see, e.g., De Roover, 1948).[3] Information does not play a role in these theories.

This paper also relates to the literature on cooperation in trade. Although this literature does not focus on the origin of banks, we can perform some interesting comparisons. Kandori (1992) and Ellison (1994) investigate a decentralized economy where agents randomly meet and play the prisoner dilemma. In their environment, cooperation can be sustained because a deviation triggers a chain of deviations - a "contagion" process - that, being the population finite, eventually backfires on the defector. In these studies institutions play no role.[4] In our environment, instead, since the

[3] Botticini and Eckstein (2005) study the involvement of Jews in mercantile and banking activities and stress the role of literacy in fostering Jews' accumulation of the skills necessary for these activities.

[4] For a recent analysis of the sustainability of cooperation in decentralized economies, see Aliprantis, Camera, and Puzzello (2007).

population is infinite, the contagion mechanism would not suffice to support cooperation; institutions are thus crucial to disseminate information and accelerate the contagion process. In Milgrom, North and Weingast (1990), fairs act as primitive courts. In fact, they are able to monitor merchants' behavior in transactions and punish non-cooperative merchants. In our decentralized environment, instead, information on transactions occurring outside institutions is private and no institution can act as a monitor.

Although Milgrom, North and Weingast (1990) are silent about banks, it is not difficult to grasp the similarities between their approach and the theory of banks as monitors of Diamond (1984). In fact, in Diamond (1984) banks emerge because they have a superior ability to monitor and punish the misbehavior of borrowers. In our environment, instead, fairs evolve towards institutions that issue tangible payment instruments, that is our emphasis is on the origin of banks as "mechanisms of settlement" in the payments system rather than as "monitoring intermediaries" between savers and borrowers. In this sense, our notion of banks is close to that in Cavalcanti and Wallace (1999) and Cavalcanti, Erosa and Temzelides (1999) in which bank notes circulate as a medium of exchange because bankers are "visible" agents. In our context, this visibility is key to prevent an overissue of notes.[5]

# 3 The Model

In this section, we first describe the environment. Then, we analyze how trade can be sustained by a "social norm" of cooperation when institutions act as fairs. Next, we analyze how trade can be supported by the circulation of tangible objects when institutions act as banks. In this first portion of the analysis, we keep the extent of trade (e.g., the population involved in trade) constant. We subsequently study how banks become essential as trade expands (e.g., the trading population grows) and formalize an endogenous transition from fairs to banks. Throughout most of the analysis, we treat the size of institutions and the process of trade expansion as exogenous. At the end of this section, we endogenize these features of the environment.

## 3.1 Environment

Time is discrete and indexed by $t \in \mathbb{N}$. The economy comprises a unit continuum of agents and a finite number $k \in \mathbb{N}$ of institutions, each inhabited by one manager. There are indivisible, perishable consumption goods. Each agent can produce one good at a cost of $c$ per unit. An agent does

[5] Dating back at least to Friedman (1959), it is well known that, in the absence of a control on the amount of notes in circulation, a likely outcome is an overissue of notes that will drive their value to zero. An extensive literature deals with the incentives not to overissue. Klein (1974) argues that agents may be encouraged to limit the issue of notes by reputational concerns (see also Araujo and Camargo, 2006); Taub (1982) and Hayek (1990) demonstrate that competition among various suppliers of notes disciplines note issuers; Ritter (1995) shows that a coalition of note issuers do not overissue because they internalize the negative effects on the value of notes. Williamson (1999) demonstrates that overissue and/or a refusal to redeem notes can be avoided when note issuers are monitored.

not derive utility from the good personally produced while she derives utility $u > c$ per unit of consumption of a good produced by another agent. If an agent stays in autarky she obtains a utility $a$ per period. Managers cannot produce but they can store goods over a period, preventing their decay. A manager derives utility equal to the measure of goods stored. Agents and managers can also costlessly issue indivisible, distinguishable objects that we label "notes". For tractability, we let each agent hold at most one note at any point in time. Agents and managers discount across periods, with $\beta \in (0,1)$ denoting their discount factor.

At the beginning of every period, each institution hosts a measure $\mu \in \left(0, \frac{1}{k}\right]$ of agents randomly drawn by nature. Inside each institution, such agents first meet the manager and then engage in one random, bilateral meeting with each other. After meetings occur inside institutions, each agent chooses whether to spend the period in autarky or engage in $n \in \mathbb{N}$ random, bilateral meetings outside institutions. To motivate the existence of trade frictions, we assume that at most one unit of a good can be produced and exchanged in a meeting (inside or outside institutions). In the event both agents in a meeting want to consume, nature randomly draws the consumer.

The crucial feature of our environment is the observability of transactions occurring inside and outside institutions: transactions in a meeting inside an institution are publicly observable while transactions in a meeting outside an institution are private, that is observable only by the agents who engage in the meeting. To fix ideas, we will often use a metaphor and imagine that institutions are located on the top of "hills" while meetings outside institutions occur in the "plain".

## 3.2 Institutions as Fairs

In this section, we show that institutions can sustain trade even if no agent or manager issues notes. They can do so because, thanks to their visibility on the top of hills, they disseminate information about agents' behavior. We simplify the description of actions by assuming that, upon entering a meeting, each agent announces "produce" or "consume". This announcement is binding within the meeting and, if the meeting involves another agent, is made before nature draws the producer.[6] We say that an agent cooperates if and only if she announces "produce". We also say that a manager cooperates if she uniformly redistributes her stored goods among the agents in her institution.

Consider the following strategy profile:

i. Managers always cooperate;

ii. An agent cooperates at the beginning of period 1. If cooperation is the only outcome she has ever observed, she participates in private exchange and cooperates inside institutions and in private meetings;

iii. If she has observed a deviation in a private meeting but no deviation inside institutions, an agent participates in private exchange but does not cooperate in private meetings or inside institutions;

iv. If she has observed deviations inside institutions, an agent stays in autarky. (a) If she has observed such deviations only once, an agent cooperates inside institutions; (b) if she has observed such deviations

[6] These assumptions do not violate an agent's participation constraint because the agent can announce "consume". At the same time, they simplify the computation of payoffs after a deviation from the equilibrium path.

more than once, an agent does not cooperate inside institutions.

Under this strategy profile, cooperation outside institutions is supported by a form of community enforcement: a deviation in a private meeting in the plain triggers a chain of deviations (see point (iii)) that, upon reaching an institution on the top of a hill, collapses all exchange in private meetings (see point (iv)), backfiring on the initial defector.[7] Unlike in Kandori (1992), where the population is finite, in our environment such a community enforcement hinges on the existence of institutions where transactions are publicly observable ("fairs"). Absent fairs, agents would have no incentive to follow the norm: in the plain contagion would occur too slowly and an agent who does not cooperate would never be reached by the contagion process and bear the consequences of her misbehavior. In contrast, fairs render the norm sustainable because they dramatically accelerate the contagion process: as soon as the non-cooperative behavior reaches a fair on the top of a hill it gets revealed to the whole economy and backfires on the initial defector. Henceforth, we say that managers and agents follow the social norm if they behave as above.

An equilibrium is defined as a strategy profile and a system of beliefs such that (i) given beliefs, the behavior of managers and agents is sequentially rational after each history and (ii) beliefs are consistent in the sense of Kreps and Wilson (1982). We prove the existence of an equilibrium in which the social norm holds. As it will become clear, our proof holds for any system of beliefs. We proceed in steps, starting with managers' incentives to follow the social norm and then turning to agents' incentives. Consider first point (i) Managers always cooperate. The proof of this point is trivial because a manager derives no benefit from keeping goods in storage for more than one period (the goods decay after that). Hence, she will always have a weak incentive to redistribute goods across the agents in her institution. Incidentally, it is worth noting here that we could render managers' incentive strict and allow them to store goods for any finite number of periods by conditioning the amount of goods a manager obtains on her past behavior. For example, we could slightly adapt point (iv) of the norm specifying that agents cooperate with managers inside institutions if and only if managers always cooperated in the past.

We now turn to agents' incentives to follow the norm on the equilibrium path, that is point ii. If cooperation is the only outcome she has ever observed, an agent participates in private exchange and cooperates inside institutions and in private meetings. We start by investigating the conditions under which an agent who has always observed cooperation has the incentive to cooperate outside institutions. Consider an agent at the beginning of her $i^{th}+1$ private meeting in period $t$, where $i \in \{0,...,n-1\}$. If she follows the social norm and cooperates, her expected (normalized) flow payoff is $\frac{1}{2}(1-\beta)(n-i)(u-c)$ and her expected (normalized) continuation payoff is

$$v_f \equiv (1-\beta)\sum_{s=0}^{\infty}\beta^s\left[k\mu+\frac{1}{2}(k\mu+n)\right](u-c), \tag{1}$$

[7] If there is a collapse of exchange in private meetings, cooperation may still continue inside fairs. This cooperation is sustained by the threat that a second deviation will be punished with lack of cooperation in any future meeting inside fairs, as in point (iii. (b)).

which, letting $\kappa \equiv k\mu$, can be rewritten as $\left[\kappa + \frac{1}{2}(\kappa + n)\right](u - c)$.[8] At the beginning of each period, with probability $\kappa$ an agent enters an institution, in which case she meets a manager, obtaining a payoff $u - c$, and another agent, obtaining in expectation a payoff $\frac{1}{2}(u - c)$ (remember that if both agents in a meeting cooperate and announce "produce" each of them has probability $\frac{1}{2}$ of being a producer and probability $\frac{1}{2}$ of being a consumer). Moreover, in each period an agent engages in $n$ private meetings, obtaining in expectation a payoff $\frac{1}{2}n(u - c)$, which henceforth we assume to exceed the flow payoff in autarky ($a$).

Suppose now that an agent deviates. In this case, her flow payoff is $(1 - \beta)(n - i)u$ and the number of agents who are involved in the contagion process triggered by her deviation up to the end of the period, including the agent herself, equals $2^{n-i}$. Therefore, there is a probability $1-(1 - \kappa)^{2^{n-i}}$ that at least one of these agents will climb a hill and enter an institution at the beginning of period $t+1$, in which case the agent's expected continuation payoff is $(1 - \beta)\frac{\kappa}{1-(1-\kappa)^{2^{n-i}}}\left[u + \frac{1}{2}(u - c) + a\right] + \beta\left[\frac{3}{2}\kappa(u - c) + a\right]$. In fact, $\frac{\kappa}{1-(1-\kappa)^{2^{n-i}}}$ is the probability that the agent will enter an institution conditional on a deviation reaching an institution. In this case, because she previously observed a deviation in a private meeting, the agent will not cooperate with the manager of the institution (obtaining $u$); thereafter, since at that point one deviation will have occurred inside an institution, the agent will cooperate with another agent inside the institution (obtaining in expectation $\frac{1}{2}(u - c)$) and then move to autarky (obtaining $a$). Moreover, using a similar reasoning, the agent's expected continuation payoff is $\beta\left[\frac{3}{2}\kappa(u - c) + a\right]$. There is a complementary probability that instead none of the $2^{n-i}$ agents reached by the contagion process will climb a hill and enter an institution in period $t + 1$ and the expected continuation payoff is $v_{2^{n-i}}$, where $v_{2^{n-i}}$ denotes the expected payoff at the beginning of private meetings when no deviation inside institutions was observed and the number of agents who are not cooperating equals $2^{n-i}$. Summing terms up and comparing, the agent will cooperate if and only if (for $i \in \{0, ..., n - 1\}$)

$$\frac{\beta}{1-\beta}\left[\frac{1}{2}n(u - c) - a\right] \geq \frac{\beta}{1-\beta}(1 - \kappa)^{2^{n-i}}(v_{2^{n-i}} - v_d) + \frac{1}{2}(n - i)(u + c) + \beta\kappa c, \tag{2}$$

where $v_d \equiv a + \beta\frac{3}{2}\kappa(u - c)$ is the expected payoff at the beginning of private meetings when a deviation inside institutions was observed once. The computation of $v_{2^{n-i}}$ is straightforward. At the beginning of a period, if an agent is not cooperating and knows that $2^{n-i} - 1$ other agents are not cooperating either, her flow payoff from following the norm is $(1 - \beta)nu$. At the beginning of the following period, since the number of agents reached by a deviation will have risen to $2^{2n-i}$, there is a probability $(1 - \kappa)^{2^{2n-i}}$ that no deviation will be observed inside institutions, in which case the expected continuation payoff is $v_{2^{2n-i}}$. There is a complementary probability that at least one of the agents who are not cooperating will enter an institution, in which case the expected continuation payoff is $(1 - \beta)\frac{\kappa}{1-(1-\kappa)^{2^{2n-i}}}\left[u + \frac{1}{2}(u - c) + a\right] + \beta\left[\frac{3}{2}\kappa(u - c) + a\right]$. Hence,

$$v_{2^{n-i}} - v_d = (1 - \beta)(nu + \beta\kappa c - a) + \beta(1 - \kappa)^{2^{2n-i}}(v_{2^{2n-i}} - v_d). \tag{3}$$

[8] At the beginning of period 1, managers have no goods in storage and the (normalized) expected payoff from following the norm is $-k\mu c + \frac{1}{2}(k\mu + n)(u - c)$.

A similar expression applies to $v_{2^{pn-i}} - v_d$, for all $p \in \mathbb{N}$. Substituting iteratively,

$$v_{2^{n-i}} - v_d = (1-\beta)(nu + \beta\kappa c - a) \sum_{s=0}^{\infty} \beta^s \left[ (1-\kappa)^{\frac{2^{n-i}}{2^n - 1}} \right]^{2^{sn}-1}. \tag{4}$$

Using the expression for $v_{2^{n-i}} - v_d$ in (4), we can then rewrite (2) as (for $i \in \{0, ..., n-1\}$)

$$\frac{\beta\left[\frac{1}{2}n(u-c) - a\right]}{1-\beta} \geq (nu + \beta\kappa c - a) \sum_{s=0}^{\infty} \beta^{s+1} (1-\kappa)^{\frac{2^{n-i}(2^{sn}+2^n-2)}{2^n-1}} + \frac{1}{2}(n-i)(u+c) + \beta\kappa c. \tag{5}$$

In the Appendix, we prove that, for any $\kappa$, there exists a unique $\beta(\kappa)$ such that condition (5) holds if and only if $\beta \geq \beta(\kappa)$.[9] In fact, an agent anticipates that a deviation in a private meeting will eventually trigger a collapse of private exchange when information on this deviation reaches an institution on the top of a hill. If she is sufficiently patient, the agent will attach a high weight to the loss of payoffs that this will entail and will have the incentive to cooperate. We also prove that, for any $\beta$, if $\kappa$ is sufficiently small the right-hand side of (5) exceeds the left-hand side and agents have no incentive to cooperate. In fact, an agent anticipates that if $\kappa$ is sufficiently small the probability that a deviation will reach an institution will be small in any period and the contagion process will be slow. Finally, observe that it is straightforward that the condition under which on the equilibrium path an agent has the incentive to cooperate inside institutions is looser than condition (5) under which an agent has the incentive to cooperate in private meetings. Intuitively, an agent knows that if she deviates inside an institution she will immediately cause a collapse of private exchange. In the Appendix, we prove this formally. This completes the demonstration of point (ii).

We now study the incentive of an agent to follow the social norm out of the equilibrium path, starting with the case in which an agent has observed deviations in private meetings in the plain. Here we provide an informal discussion while we relegate formal details to the Appendix. iii. If she has observed a deviation in a private meeting but no deviation inside institutions, an agent participates in private exchange but does not cooperate in private meetings or inside institutions. After observing a deviation in a private meeting, an agent may be tempted to cooperate in order to slow down the spread of non-cooperative behavior and reduce the probability that a deviation reaches an institution. The agent is not tempted to do so if and only if the probability that anyway other agents will spread the non-cooperative behavior is sufficiently high. In turn, the probability that other agents deviate inside institutions is higher the larger the measure of agents who climb hills and enter institutions, i.e. the higher is $\kappa$. Therefore, for a given $\beta$, an agent is not tempted to cooperate if and only if $\kappa$ is sufficiently high. Finally, we also need to prove point iv. (a) If she has observed deviations inside institutions only once, an agent stays in autarky but cooperates inside institutions; (b) if she has observed deviations inside institutions more than once, an agent stays in autarky and does not cooperate inside institutions. The proof of this point is also presented in the Appendix. We obtain again that, for a given $\beta$, $\kappa$ cannot be too low, that is the size of institutions relative to the trading population cannot be too small.

[9] We stress the dependence of parameter thresholds on $\kappa$ and $\beta$. However, the reader should keep in mind that the other parameters ($n$, $u$, $c$, and $a$) matter as well.

Proposition 1 wraps up and summarizes our first result.

**Proposition 1** *Choose a vector of parameters* $(\beta, \kappa, a, c, n, u) \equiv (\beta, \kappa, .)$. *There exists a discount factor* $\beta^* (\kappa, .) \in (0, 1)$ *and a probability that an agent enters an institution* $\kappa^* (\beta, .) \in (0, 1)$ *such that the social norm is an equilibrium if and only if the chosen values of* $\beta$ *and* $\kappa$ *jointly satisfy* $\beta \geq \beta^* (\kappa, .)$ *and* $\kappa \geq \kappa^* (\beta, .)$. *In particular,* $\lim_{\beta \to 1} \kappa^* (\beta, .) < 1$ *and the social norm is an equilibrium for all* $\beta \geq \beta^* (\kappa, .)$ *and* $\kappa \geq \lim_{\beta \to 1} \kappa^* (\beta, .)$.

**Proof.** In the Appendix. ■

For our purposes, the most important message of the proposition is that fair-based trade can only occur if the probability $\kappa$ that an agent enters an institution is not too small. This result will play a critical role in the remainder of the analysis.

### 3.3 Institutions as Banks

In this section, we demonstrate that institutions can also sustain trade by allowing the use of notes as a payment instrument. In our economy, transactions in private meetings in the plain are not publicly observable and the population is large. Therefore, if notes are valued, an agent may have no incentive to produce in exchange for a note in a private meeting because she can always issue a note herself. If all agents behave this way, no production will occur in exchange for notes in private meetings and, absent other mechanisms of trade, at the beginning of every period all agents will choose to stay in autarky. The same is not true if the managers of institutions issue notes. In fact, since transactions occurring inside institutions are publicly observable, managers can be immediately punished by agents if they overissue notes.

The analysis in this section builds on Cavalcanti and Wallace (1999).[10] When notes circulate, our environment is no longer static because the amount of notes can change across periods. Here we focus on a steady state equilibrium while in the Appendix we analyze a transition to this steady state. The behavior of agents and managers can be described by agents' decision whether to participate in private exchange and by the vector $\mathsf{V} = (\lambda_1, \lambda_0, \gamma_1, \gamma_0, \alpha_1, \alpha_0)$, where $\lambda_1$ is the probability that an agent produces in exchange for a note (inside or outside an institution), $\lambda_0$ is the probability that an agent produces if she is not paid with a note (inside or outside an institution), $\gamma_1$ is the probability that inside an institution a manager transfers one unit of good to an agent with a note (i.e. redeems notes), $\gamma_0$ is the probability that a manager transfers goods to agents without notes, $\alpha_1 \in [0, 1]$ is the probability that a manager gives a note to an agent who has just redeemed a note, and $\alpha_0 \in [0, 1]$ is the probability that a manager gives a note to an agent who has not redeemed a note. In the Appendix, we prove that there exists a steady state equilibrium in which agents always participate in private exchange and $\mathsf{V} = (1, 0, 1, 0, 1, 0)$, i.e.: (i) agents always produce in exchange for notes

[10] Note that in Cavalcanti and Wallace (1999) bankers are agents while in our study banks are institutions (or locations).

$(\lambda_1 = 1)$ but never produce if they are not paid with notes $(\lambda_0 = 0)$, (ii) inside institutions managers always redeem notes $(\gamma_1 = 1)$ while they do not transfer goods to agents without notes $(\gamma_0 = 0)$, (iii) managers give notes in exchange for goods to all agents who have just redeemed a note, and only to them ($\alpha_1 = 1$ and $\alpha_0 = 0$). Moreover, if in any period a manager deviates from the behavior described by the vector $\mathsf{V} = (1, 0, 1, 0, 1, 0)$ no agent produces from that period on.

A steady state equilibrium is defined not only by agents' and managers' behavior but also by the measure $m$ of agents with notes and by agents' and managers' value functions. It is useful to display the value functions of agents that the candidate equilibrium vector $\mathsf{V} = (1, 0, 1, 0, 1, 0)$ implies for a given $m$.[11] Let $v_j^i(m)$ denote the expected payoff of an agent with $j$ notes right before her $i^{th}$ meeting with another agent, where $j \in \{0, 1\}$ and $i \in \{0, ..., n\}$. We let $i = 0$ correspond to the meeting with another agent inside an institution. Moreover, let $v_j^{n+1}(m)$ denote the expected payoff of an agent with $j$ notes at the end of her private meetings in a period. For all $i \in \{0, ..., n\}$,

$$
\begin{aligned}
v_1^i(m) &= m v_1^{i+1}(m) + (1-m)\left[u + v_0^{i+1}(m)\right], && (6)\\
v_0^i(m) &= m\left[-c + v_1^{i+1}(m)\right] + (1-m) v_0^{i+1}(m), && (7)\\
v_1^{n+1}(m) &= \beta\left\{\kappa\left[u - c + v_1^0(m)\right] + (1-\kappa) v_1^1(m)\right\}, && (8)\\
v_0^{n+1}(m) &= \beta\left[\kappa v_0^0(m) + (1-\kappa) v_0^1(m)\right]. && (9)
\end{aligned}
$$

Consider $v_1^i(m)$ (a similar interpretation holds for $v_0^i(m)$). An agent with a note right before her $i^{th}$ private meeting has probability $m$ of meeting another agent with a note. In this case, no exchange occurs because neither agent wants to produce for a note given the upper bound on note holdings. With probability $1 - m$, instead, the agent meets an agent without a note, pays one unit of a good with her note, obtaining utility $u$, and moves to the next meeting without a note. Consider next $v_1^{n+1}(m)$. At the end of a period, an agent with a note expects to climb a hill and enter an institution with probability $\kappa$ at the beginning of the subsequent period, in which case she will redeem her note in exchange for one unit of good (obtaining $u$), will produce for a manager (at the cost $c$) and will obtain a new note (with the complementary probability the agent will not enter an institution). Finally, consider $v_0^{n+1}(m)$. At the end of a period, an agent without a note expects to enter an institution with probability $\kappa$, in which case she will not obtain notes or goods (with the complementary probability the agent will not enter an institution).

Proposition 2 establishes our second result, that is conditions under which the candidate steady state equilibrium with bank-based trade exists.

**Proposition 2** *Choose a vector of parameters ($\beta$,$\kappa$,$a$,$c$,$n$,$u$). If*

$$\beta \geq \frac{c}{c + (1-m)(u-c)} \tag{10}$$

*and*

$$nm(1-m)(u-c) \geq a + (1-\beta) m \left[(1-m) u + mc\right], \tag{11}$$

[11] The value function of a manager is simply $v^M = \mu m u$.

*there exists a steady state equilibrium in which agents participate in private exchange, agents and managers behave according to* $\mathsf{V} = (1, 0, 1, 0, 1, 0)$*, and the measure of agents with notes equals* $m$*.*

**Proof.** In the Appendix. ■

We have already discussed how the visibility of institutions on the top of hills prevents managers from overissuing notes. Proposition 2 demonstrates that notes issued by managers can circulate as long as $\beta$ is sufficiently high and in each period the frequency of meetings in which exchange occurs (as measured by $nm(1-m)$) is sufficiently large. In particular, inspection of conditions (10) and (11) in the proposition reveals a critical property of the bank-based equilibrium. Unlike in the case of fair-based trade, as long as $\beta$ is sufficiently high, such an equilibrium exists regardless of the size of institutions relative to the trading population, as captured by $\kappa$. The economic intuition is simple. The value of notes has two components: their value as a payment instrument and their redemption value. Even though the redemption value is close to zero when the relative size of institutions is small (because agents expect to enter institutions and be able to redeem notes with a low probability), this can always be offset by a sufficiently high value as a payment instrument. This, in turn, can be attained by appropriately setting the amount of notes in circulation and, hence, the frequency of meetings in which exchange occurs. Put differently, even if institutions are sparse and, hence, of little value for agents in redeeming notes, they can preserve an essential role by costlessly issuing notes that serve as a payment instrument (i.e. providing liquidity to the economy). Furthermore, regardless of how large the population involved in trade is relative to the size of institutions, only one piece of information - the amount of notes - is needed to prevent an overissue of notes.

From a welfare point of view, it is important to observe that the "bank-based equilibrium" in Proposition 2 is Pareto dominated by the "fair-based equilibrium" in Proposition 1. The reason is that the social norm allows transactions to occur in all private meetings whereas note-exchange does not because it requires that in a meeting one agent has a note and her match has zero notes.

## 3.4 The Evolution of Institutions

In the previous two sections, we have established conditions under which fair-based trade and bank-based trade can occur in equilibrium. In this section, we let the extent of trade change over time and, building on the results of Propositions 1 and 2, we demonstrate that when trade expands institutions must eventually perform as banks in order for exchange to occur in private meetings (that is, in the long run banks are "essential"). We also construct an equilibrium under which the economy endogenously transits from fair-based trade to bank-based trade. We let the expansion of trade be driven by an increase in the measure of agents participating in the exchange process. Precisely, we slightly modify the previous environment and assume that, on top of the unit measure of agents initially populating the economy, at the very beginning of each period $t \in \{2, ..., \bar{t}\}$ ($t > \bar{t}$) a measure $\eta$ (zero) of agents is born, implying that the measure of agents in $t \leq \bar{t}$ ($t > \bar{t}$) equals $\eta_t = 1 + \eta(t-1)$ (respectively, $\eta_t = 1 + \eta(\bar{t}-1)$). In every period, after new agents are born, institutions host a measure $\kappa$ of agents randomly chosen by nature. Hence, the probability $\frac{\kappa}{\eta_t}$ that in period $t$ an agent climbs a hill and enters an institution decreases up to period $\bar{t}$ and takes on the value of $\frac{\kappa}{1+\eta(\bar{t}-1)}$ from period $\bar{t}$ onwards.

**Essentiality of Banks.** Proposition 2 implies that the drop in the probability that agents enter institutions induced by population growth poses no problem for bank-based trade because managers can always inject additional notes into the economy to offset its impact. In contrast, Proposition 1 implies that the drop in this probability may lead to the collapse of fair-based trade. Building on these two results, in Proposition 3 we prove that, if trade expands without bounds, in the long run exchange outside institutions can only be sustained if institutions act as banks.

**Proposition 3** *Assume that the trading population grows without bounds, that is $\bar{t} \equiv \infty$. In the long run, banks are essential, that is exchange outside institutions can only occur if managers issue notes.*

**Proof.** Assume that there exists an equilibrium where in the long run exchange outside institutions does not involve the use of notes. Consider the incentive of an agent to announce produce in the $i^{th}$ private meeting ($i \in \{1, ...n\}$) in some period $t$. If the agent announces produce, her net expected current payoff is $-\frac{u+c}{2}$. In order to be willing to incur this cost, she must receive a net continuation payoff at least equal to $\frac{1}{2}c$. Since the economy is populated by a continuum of agents, this can only occur if the information of the announcement made in the $i^{th}$ private meeting in period $t$ reaches an institution in some future period. Now, we can always choose $t$ sufficiently large that the probability that this information reaches an institution is sufficiently small. As a result, in any equilibrium that does not involve the use of notes, the agent has no incentive to announce "produce" outside institutions when $t$ is sufficiently large. Finally, as shown in Proposition 2, as long as agents are sufficiently patient and the frequency of meetings in which exchange occurs is large enough, there exists an equilibrium where notes issued by managers are valued as a medium of exchange. This is true irrespective of the relative size of institutions. ■

Proposition 3 considers the case in which $\bar{t} \equiv \infty$. Observe, however, that its argument also applies when $\bar{t} < \infty$. In this case, a straightforward modification of the proof implies that, for any discount factor $\beta$, there always exists $\bar{t}(\beta)$ sufficiently large that banks are essential for all $\bar{t} \geq \bar{t}(\beta)$.

**Formalizing a Transition.** In the remainder of this section, we demonstrate with an example that when trade expands the economy can endogenously transit from a state in which institutions act as fairs to a state in which institutions act as banks. To show this, we introduce a second modification to our environment, besides population growth, allowing for heterogeneity in agents' production costs. As we elaborate below, agents' production cost affects their willingness to follow the social norm and, hence, exerts a critical role for the sustainability of the norm. We assume that with probability $1 - \theta$ in every period all agents have production cost $c$. There is also a probability

$\theta(1-\varepsilon)$ $(\theta\varepsilon)$ that a state $\mathcal{S}_1$ $(\mathcal{S}_2)$ is realized in which an agent (respectively, a share $\pi$ of the agents) already populating the economy in period 1 has (have) production cost $c_H > c$, while all other agents have production cost $c$. We treat states $\mathcal{S}_1$ and $\mathcal{S}_2$ as perturbations to the distribution of production costs, working with the case of an economy where $\theta$ and $\varepsilon$ are close to zero. Finally, the production cost of an agent is her private information. By modifying the environment this way, we will be able to construct an equilibrium where, regardless of her production cost, each agent faces a trade-off between ensuring the continuation of fair-based trade and triggering an eventual transition to bank-based trade. Such an equilibrium has two desirable features. First, it allows for a transition from fair-based to bank-based trade that is not triggered by coordination reasons, where an agent deviates in a period simply because she knows that another agent deviates in the same period. Second, it prevents the unravelling of a transition, where an agent deviates in a period simply because she anticipates that another agent (maybe herself) will deviate in the subsequent period.

The reasoning that follows relies on Section 3.2. We capture the transition from fairs to banks by retaining points (ii) and (iii) but appropriately adapting points (i) and (iv) of the social norm, that is the points that describe, respectively, the behavior of managers (point (i)) and what happens after deviations are observed inside institutions (point (iv)). Consider first point (i). While in the framework of Section 3.2 managers always cooperate, we now propose a strategy such that managers cooperate after histories of cooperation inside institutions and behave as in Proposition 2 otherwise, issuing and redeeming notes. Consider next point (iv). In the framework of Section 3.2, after observing a deviation inside institutions, an agent stays in autarky in the current and in all future periods. We now propose a strategy such that, after observing such a deviation, an agent switches to the behavior in Proposition 2 and only produces in exchange for notes thereafter. Specifically,

i.(adapted) If cooperation is the only outcome she has ever observed, a manager cooperates. If she observes deviations inside institutions, in the period in which the deviation occurs a manager distributes notes to a fraction m of agents inside her institution; in all following periods, she (a) transfers one unit of good to any agent with a note (redeems notes) while she does not transfer goods to agents without notes; (b) gives notes in exchange for goods to all agents who have just redeemed a note, and only to them;

iv.(adapted) If she has observed deviations inside institutions, an agent always produces in exchange for a note but she never produces if she is not paid with notes.

In Proposition 1, we have proved that on the equilibrium path an agent has the incentive to follow the social norm if and only if (for all $i \in \{0, ..., n-1\}$)

$$\frac{\beta}{1-\beta}\left[(v_f - v_d) - (1-\kappa)^{2^{n-i}}(v_{2^{n-i}} - v_d)\right] \geq \frac{1}{2}(n-i)(u+c) + \frac{\beta\kappa}{2}(3u-c). \tag{12}$$

On the basis of the discussion above, we need to modify condition (12) in four ways. The first three simply stem from the fact that now the trading population grows over time; the fourth stems from the adaptation of points (i) and (iv) of the norm. First of all, in period $t$ the probability that an agent climbs a hill and enters an institution is $\frac{\kappa}{\eta_t}$ and no longer $\kappa$. Second, consider $v_f$, that is the agent's expected payoff from following the norm. Since the probability $\frac{\kappa}{\eta_t}$ that an agent enters an institution changes over time, $v_f$ is replaced by

$$\widetilde{v}_{f,t} = \frac{1}{2} n \left(u - \widetilde{c}\right) + \frac{3}{2} \left(u - \widetilde{c}\right) \sum_{s=0}^{\infty} \left(1 - \beta\right) \beta^s \frac{\kappa}{\eta_{t+s}}, \tag{13}$$

where $\widetilde{c} \in \{c, c_H\}$. Third, consider $v_{2^{n-i}}$, that is the expected payoff at the beginning of private meetings when no deviation was observed inside institutions and the number of agents who are not cooperating equals $2^{n-i}$. Since the probability $\frac{\kappa}{\eta_t}$ that an agent enters an institution changes over time, this payoff also changes over time and is replaced by $\widetilde{v}_{2^{n-i},t}$. Fourth, consider $v_d$, that is the expected payoff at the beginning of private meetings right after a deviation is observed inside institutions. In the framework of Section 3.2, $v_d \equiv a + \beta \frac{3}{2} \kappa \left(u - c\right)$. With the adaptation of points (i) and (iv) of the norm, $v_d$ is replaced by $\widetilde{w}_t\left(m\right)$, where $\widetilde{w}_t\left(m\right)$ is the expected payoff in period $t$ when notes circulate and the measure of agents with notes is $m$.

Taking these four changes into account, on the equilibrium path an agent with production cost $\widetilde{c}$ has the incentive to follow the adapted social norm in period $t$ if and only if (for all $i \in \{0, ..., n-1\}$)

$$\frac{\beta}{1-\beta} \left\{ \begin{array}{c} \left[\widetilde{v}_{f,t+1} - \widetilde{w}_{t+1}\left(m\right)\right] - \\ \left(1 - \frac{\kappa}{\eta_{t+1}}\right)^{2^{n-i}} \left[\widetilde{v}_{2^{n-i},t+1} - \widetilde{w}_{t+1}\left(m\right)\right] \end{array} \right\} \geq \left[ \begin{array}{c} \frac{1}{2}\left(n-i\right)\left(u + \widetilde{c}\right) + \\ \frac{\beta\kappa}{2\eta_{t+1}}\left(3u - \widetilde{c}\right) \end{array} \right]. \tag{14}$$

The computation of $\widetilde{v}_{2^{n-i},t+1}$ is analogous to that of $v_{2^{n-i}}$ in Section 3.2. If an agent is not cooperating and knows that $2^{n-i} - 1$ other agents are not cooperating either, her flow payoff from following the norm is $\left(1-\beta\right) nu$. At the beginning of the following period, there is a probability $\left(1 - \frac{\kappa}{\eta_{t+2}}\right)^{2^{2n-i}}$ that a deviation will not be observed inside institutions, in which case her expected continuation payoff is $\widetilde{v}_{2^{2n-i},t+2}$. There is a complementary probability that at least one of the agents who are not cooperating will climb a hill and enter an institution, in which case her expected continuation payoff is $\left(1-\beta\right) \frac{\frac{\kappa}{\eta_{t+2}}}{1 - \left(1 - \frac{\kappa}{\eta_{t+2}}\right)^{2^{2n-i}}} u + \widetilde{w}_{t+2}\left(m\right)$. Thus,[12]

$$\widetilde{v}_{2^{n-i},t+1} - \widetilde{w}_{t+1}\left(m\right) = \left(1-\beta\right) \sum_{s=0}^{\infty} \beta^s \prod_{j=2}^{s+1} \left(1 - \frac{\kappa}{\eta_{t+1+j}}\right)^{2^{jn-i}} \left[nu - \widetilde{w}_{t+s+2}\left(m\right) + \frac{\beta\kappa}{\eta_{t+s+3}} u\right]. \tag{15}$$

Using (15), condition (14) can then be rewritten as (for all $i \in \{0, ..., n-1\}$)

$$\frac{\beta}{1-\beta} \left[\widetilde{v}_{f,t+1} - \widetilde{w}_{t+1}\left(m\right)\right] \geq \left\{ \begin{array}{c} \sum_{s=0}^{\infty} \beta^{s+1} \prod_{j=1}^{s+1} \left(1 - \frac{\kappa}{\eta_{t+j}}\right)^{2^{jn-i}} \left[nu - \widetilde{w}_{t+s+2}\left(m\right) + \frac{\beta\kappa}{\eta_{t+s+2}} u\right] + \\ + \frac{1}{2}\left(n-i\right)\left(u+\widetilde{c}\right) + \frac{\beta\kappa}{2\eta_{t+1}}\left(3u - \widetilde{c}\right) \end{array} \right\}. \tag{16}$$

[12] We define $\prod_{j=2}^{1} \left(1 - \frac{\kappa}{\eta_{t+j}}\right)^{2^{jn-i}} \equiv 1$.

To grasp the intuition behind the collapse of fair-based trade and the transition to bank-based trade it is now sufficient to inspect condition (16). Start with observing that the left hand side of (16) is strictly positive. In fact, under fair-based trade production and consumption occur in all meetings, while under bank-based trade production in a meeting requires that one agent has a note while her match has zero notes. Therefore, the expected payoff on the equilibrium path under fair-based trade is strictly higher than the expected payoff on the equilibrium path under bank-based trade, that is $\widetilde{v}_{f,t+1} > \widetilde{w}_{t+1}(m)$. Next, inspection of (16) reveals that for any $\beta$ and $\widetilde{c}$ there exists a sufficiently large measure $\eta_{t(\beta,\widetilde{c})}$ of the trading population (a sufficiently small probability $\frac{\kappa}{\eta_{t(\beta,\widetilde{c})}}$ that an agent enters an institution) that condition (16) holds for $\eta_t < \eta_{t(\beta,\widetilde{c})}$ while it is violated for $\eta_t \geq \eta_{t(\beta,\widetilde{c})}$. In particular, the difference between the left and the right hand side of (16) and, hence, $\eta_{t(\beta,\widetilde{c})}$ is strictly decreasing in the production cost $\widetilde{c}$ of the agent. The intuition is fairly simple. The higher her production cost, the larger is the short-run gain of the agent if she stops cooperating and continues to consume without producing until a deviation reaches a fair and is publicly disclosed. Furthermore, the difference between her expected continuation payoff from fair-based trade and her expected continuation payoff from bank-based trade is strictly decreasing in her production cost.[13]

Using the reasoning above, we are now in a position to write down this result.

**Proposition 4** *Assume that the state $\mathcal{S}_2$ is realized in which a positive measure of agents have production cost $c_H$, while the complement has cost $c$. There exists a region of the parameters $\beta$,$\kappa$,$a$,$c_H$,$c$, $n$,$u$,$\eta$ and $\bar{t}$ such that up to some period $t(\beta, c_H) \leq \bar{t}$ all agents cooperate while at $t(\beta, c_H)$ all $c_H$ agents stop cooperating and, as a result, the economy eventually transits from fair-based trade to bank-based trade.*

**Proof.** In the Appendix. ∎

Proposition 4 completes our core analysis. The proposition integrates Proposition 3, showing that not only in the long run bank-based trade can be the only viable form of exchange when trade expands but also that the transition from fair-based trade to bank-based trade can occur as an equilibrium outcome.

## 3.5 Robustness Issues

We conclude the model by exploring robustness issues. We are especially interested in evaluating how some features of the environment, such as the process of trade expansion and the size of institutions, can be endogenized. At the end of this section, we also discuss how the results are robust to modifying ancillary assumptions of the model.

[13] Formally, the left hand side of (16) is decreasing in $\widetilde{c}$. Let us now show that the right hand side of (16) is increasing in $\widetilde{c}$. First of all, $\widetilde{w}_{t+s+2}(m)$ is decreasing in $\widetilde{c}$, hence $-\widetilde{w}_{t+s+2}(m)$ is increasing in $\widetilde{c}$. Moreover, the last two terms in the curly brackets can be rewritten as $u\left[\frac{1}{2}(n-i)+\frac{\beta\kappa}{2\eta_{t+1}}3\right]+\widetilde{c}\left[\frac{1}{2}(n-i)-\frac{\beta\kappa}{2\eta_{t+1}}\right]$. The term that multiplies $\widetilde{c}$ is always positive.

**Endogenizing the Expansion of Trade.** In the model, we have assumed that the expansion of trade is generated by an exogenous process of population growth. However, our environment offers a natural way to endogenize the increase in the extent of trade which builds on the distinctive feature of institutions (their visibility). An interesting side result of this variation of the model is that it highlights the possibility that, by promoting the expansion of trade through their visibility, fairs themselves create the conditions for their own demise. Formally, consider the following variation. There is a probability $1-\theta$ that in all periods a share $\pi_{HH}$ of newborn agents have production cost $c_{HH} >> c$ and a share $1-\pi_{HH}$ have production cost $c$. With probability $\theta(1-\varepsilon)$ $(\theta\varepsilon)$, a state $\mathcal{S}_1$ $(\mathcal{S}_2)$ is realized in which a share $\pi_{HH}$ of the agents populating the economy in period 1 have production cost $c_{HH}$, a randomly chosen agent (a share $\pi_H$ of these agents) has (have) production cost $c_H \in (c, c_{HH})$, and the complement has production cost $c$. In all other periods, a share $\pi_{HH}$ of newborn agents have production cost $c_{HH}$ and a share $1-\pi_{HH}$ have production cost $c$. As before, we work with the limit case of an economy where $\theta$ and $\varepsilon$ are close to 0.

In this modified environment, if $c_{HH}$ is sufficiently high no agent with such a production cost will be willing to produce. It is straightforward to show that, in turn, as long as $\pi_{HH} > \frac{u-c}{u+c}$ an agent is unwilling to engage in a private meeting with another agent if she cannot discern whether her match would be willing to produce or not. Since transactions inside institutions are publicly observable, an agent with production cost $c$ or $c_H$ can reveal that she is not of the $c_{HH}$ type by producing for the manager of the institution she enters.[14] This makes clear the crucial role that the visibility of institutions on the top of hills plays in integrating traders into the exchange process: institutions can act as "gates" through which agents gain access to trade. The rest of the analysis unfolds as in the main version of the model.[15]

**Endogenizing the Size of Institutions.** In the model, we have assumed that the measure of agents that institutions host is fixed. As a result, the growth of the trading population leads to a drop in the probability that an agent enters an institution. The reader may argue, however, that the size of fairs could also increase over time and this could compensate for the expansion of trade. In a sense, this could be interpreted as fairs providing additional records and information to the economy in a similar way as banks do when injecting additional notes. However, a key difference between increasing the number of notes and increasing the size of fairs is that the latter is costly, while notes are costless to produce. In what follows, we formalize this intuition.

Consider a modified environment where, for simplicity, there is only one institution ($k=1$) and assume that at the beginning of period 1 the manager of the institution chooses its size - that is the measure $\mu_t$ of agents the institution can accommodate - for any period $t$. If in a period the size of the institution increases by $\mu_t - \mu_{t-1}$ the manager will suffer an extra disutility $\phi(\mu_t) - \phi(\mu_{t-1})$ from maintaining this additional size. We let $\phi(\mu_t)$ be a standard strictly increasing and strictly convex

[14] We are implicitly assuming that, once a $c$ or $c_H$ agent has revealed her type, she can always be recognized as such in her private meetings.

[15] Details are available upon request.

neoclassical cost function. Let us now turn to the utility gain the manager derives from a size change. This can comprise two components. The first is simply the extra flow utility that, conditional on trade being fair-based or bank-based, the manager obtains from storing a larger amount of goods. This equals the extra measure $\mu_t - \mu_{t-1}$ ($m(\mu_t - \mu_{t-1})$) of goods the manager obtains from agents under fair-based (bank-based) trade. The second component reflects instead the impact that the size change can have on agents' behavior and, hence, on the transition to bank-based trade. In fact, a larger size of the fair implies a higher probability that an agent enters the fair, which in turn tends to prevent the collapse of fair-based trade. This raises the manager's utility because, for a given size $\mu_t$ of the institution, the flow measure of goods she obtains under bank-based trade ($m\mu_t$) is smaller than under fair-base trade ($\mu_t$) - remember that under bank-based trade the manager must give notes to obtain goods and she cannot give notes to everyone.

Consider now the adapted social norm in Section 3.4 and revisit Proposition 4. The manager's decision problem is simple. The sequence of $\{\mu_t\}$ that maximizes her expected utility depends on her expectation regarding agents' behavior. For instance, in the region of the parameter space identified in Proposition 4, if all agents have production cost $c$ and the manager is aware of that, even if the manager chooses to never increase the size of the institution, all agents will always cooperate. As a result, the only benefit of increasing the size of the institution will consist of the increase in the measure of goods that the manager obtains from agents under fair-based trade. This is no longer true if some agents have production cost $c_H$. For example, if the state $\mathcal{S}_2$ is realized, and the manager is aware of that, the manager will derive two benefits from increasing the size of the institution. First, this will raise the measure of goods that she obtains from agents under both fair-based and bank-based trade. Second, if the increase is large enough, this will prevent the deviation of $c_H$ agents and, hence, the transition from fair-based trade to bank-based trade. It is straightforward to show, however, that, as long as the disutility $\phi(\mu_t)$ of maintaining the size of the institution grows sufficiently fast, the manager will choose not to increase the size of the institution enough to prevent the transition.

**Further Issues.** In the model, we have introduced two technical assumptions that deserve further scrutiny. First, the assumption of a continuum of agents is made because under the social norm it simplifies the computation of an agent's expected payoff following a deviation from the equilibrium path. Precisely, it ensures that the probability of a meeting between two agents who have observed a deviation equals zero, so that there is no need to take such meetings into account when computing the number of agents who are reached by a deviation (thus, the probability that a deviation reaches a fair) at the end of every period. Note that this assumption does not drive our results. Indeed, we could work with a finite population and show that, for any discount factor, the social norm cannot be sustained if there are no institutions and the population is sufficiently large.

A second assumption we have made is that, besides agents, managers populate the economy. This assumption allows us to endogenize the issue of notes. An alternative approach would be to assume that, when institutions act as banks, nature randomly selects some agents to act as bankers. Because there is a finite number of institutions, and hence a zero probability that any particular agent is selected, an agent would not take the probability of becoming a banker into account. The drawback of this alternative specification is that it would generate a technical problem in the computation of agents' expected utility. In fact, because managers meet a continuum of agents, we have assumed that an agent's utility is a function of the number of goods she consumes while a manager's utility depends on the measure - not on the number - of goods she stores.

# 4 Some Historical Facts

In this section, we evaluate our hypothesis on the origin of banks in light of the historical experience of Medieval and Early Modern Europe. Our objective is not to obtain hard evidence but simply to verify whether the historical facts are broadly consistent with our hypothesis. We stress "broadly" because this long period featured a wide variety of financial institutions and instruments. We first evaluate whether our key assumption - the public visibility of fairs and banks - is grounded in the account of historians. Next, we discuss the more traditional view of fairs as "primitive courts" that could monitor and punish merchants. Finally, we turn to historical evidence on the transition from fair-based trade to bank-based trade. Specifically, we evaluate whether banks indeed emerged as the evolution of fairs and whether this evolution stemmed from the inability of fairs to keep up with trade expansion. The Cambridge Economic History of Europe is the main source of the historical evidence presented below.

**Public Information on Transactions in Fairs and Banks.** There is historical evidence that a sizable volume of public information and records was available about transactions occurring in Medieval fairs. Verlinden (1965) reports:

> "It is therefore certain that there existed at the Champagne fairs a real records department (p. 128). [...] [The fairs] of Ypres gave rise to the compilation of important series - more than 7,000 specimens - of registered obligations (*lettres de foire*) of which the oldest examples, destroyed in 1914, dated back to the middle of the thirteenth century (p. 137)."

Similarly, there is also historical evidence on the large amount of public information that was available about transactions occurring in banks in Medieval and Early Modern Europe. In particular, the books of banks were publicly available for inspections and, hence, transactions could easily be detected. De Roover (1948, pp. 265-266) reports that

> "Bank journals were considered in most Italian cities as public [...] records. [...] The journal had to be kept strictly according to chronological order, without blanks and without erasures. [...] [In fact] the public character of bank records made such a formality [to require a voucher for each deposit] superfluous."

**Fairs and Legal Enforcement.** The literature (see, e.g., Milgrom, North and Weingast, 1990) has sometimes interpreted fairs as primitive courts whose main role would have been to monitor merchants and punish them whenever they did not comply with contractual obligations. The account of historians suggests that this can only capture one aspect of the role of fairs. In fact, the ability to enforce promises was very limited in Medieval and Early Modern Europe:

> "Independent merchants, unfortunately, were entirely at the mercy of the correspondents to whom they sent goods on consignment. Usually there was no remedy against agents who were ill-chosen and proved to be either inefficient or dishonest (De Roover, 1965, p. 87)."

Even when fairs carried out some form of legal enforcement, this practice took time to spread and was often summary. According to Verlinden (1965),

> "The wardens [of fairs] are attested from 1174 but more than a century was to elapse before they acquired the higher jurisdiction which characterized them at the end of the thirteenth century (p. 131). [...] in Flanders [the exercise of justice and the procedure in their courts] was usually summary (pp. 136-37)."

**The Evolution of Fairs towards Financial Institutions.** The progressive transition from fairs to banks is well documented by historians. Verlinden (1965, pp. 132-33) writes:

> "In any case, from this time [beginning of the XIV century] onwards the fairs began to lose much of their truly commercial importance. [...] The chief function of the fairs now became the regulation of the capital market. [...] In fact the financial system of the Lyons fairs in the fifteenth century, the Besançon fairs in the sixteenth and those of Piacenza up to the seventeenth, was already heralded in Champagne by the end of the thirteenth century [...]. If however the fairs are considered from the point of view of their commercial character it seems that a recession set in from about 1260; but as financial markets they still enjoyed a considerable boom and this prosperity continued until about 1320. After this date their decline was pronounced and the last important group of Italians, that of Piacenza, disappeared in 1350. This group, however, was composed of financiers."

The so called "fair banks" offer an illuminating example of the transition. Indeed, this hybrid of fairs and banks were allegedly at the root of the development of banking in several parts of Europe. Van der Wee (1977) well describes this:

> "At the Castilian fairs a special combination between clearing bank system and fair payments was obtained through the creation of fair banks (*bancos de feria*) (p. 316). [...] [At the Lyons fairs] first of all there were the quarterly commercial fairs, each lasting about a fortnight. Then there were the fairs of payment, which lasted about one week. [The latter] gradually acquired an independent status, developing into specific *foires de change* or fairs of exchange, similar to those of Castile (p. 320). [Moreover] The *foires de change* of Besançon and Piacenza under the direction of the Genoese merchant bankers formed the apogee in the institutional development of banking and credit in southern Europe from the sixteenth to the beginning of the seventeenth century (p. 320)."

The historical evidence also suggests that the expansion of trade was one of the main determinants of the transition from fairs to banks. In this regard, De Roover (1965) writes:

> "This situation was greatly altered as the fairs declined [...] As a result [of trade expansion], it became more difficult for the merchant to keep track of the customs of the different places of traffic (p. 94)."

In particular, De Roover (1965) argues that one of the main problems faced by fairs was that, as a result of the growing volume of trade, they progressively became unable to efficiently record all transactions. Indeed, looking at the activity of the notaries who recorded transactions in fairs,

> "It was inconvenient and time-consuming to approach a notary for every business transaction of any importance. This inconvenience was felt more and more as the volume of business grew (p. 69)."

# 5 Discussion

Before concluding our analysis, it is worth discussing more in detail the mechanisms of exchange under fair-based trade and bank-based trade. In this section, we first explore a possible reinterpretation of the mechanism of multilateral reciprocity under fair-based trade. We next consider alternative interpretations of the tangible payment instrument used under bank-based trade.

**Fair-Based Trade: Multilateral Reciprocity or Credit?** In the model, we have considered a somewhat extreme notion of cooperation under fair-based trade that resembles anonymous gift-exchange: an agent produces for another agent not because she expects that the latter will produce for her in the future but because she expects that other agents will do so. An opposite scenario would be one in which agents cooperate one with another because they expect to regularly meet again in the future. In this case, standard repeated interaction effects could sustain cooperation without the need of institutions. Our model thus stresses the relevance of institutions in a context where agents do not expect to regularly meet over time. Note that we could modify our environment and, for example, allow two agents to meet twice (or also a finite number of times). In a first meeting an agent would produce for the other as long as she expects that in the second meeting her match will reciprocate. This specification would resemble credit, with the production in the second meeting constituting the repayment. In this specification, cooperating would thus mean "repaying" in the second meeting while deviating would mean "defaulting". The analysis would carry through to this modified environment.

**Bank-Based Trade: Interpreting the Payment Instrument.** In the model, building for example on Cavalcanti and Wallace (1999) and Cavalcanti, Erosa and Temzelides (1999), we have focused on an environment where banks issue tangible objects that circulate as media of exchange. It is well known that banks have played this role in several historical periods (e.g., in the U.S. during the 19th century).[16] Bankers started playing this role in Early Modern Europe. In The Cambridge Economic History of Europe, Van der Wee (1977, p. 314) writes:

> "The Italian public banks also created money. The Neapolitan public banks, for example, entered the deposits in a *madrefede*, and on the strength of this the clients supplied their creditors with cheques (*polizze*) drawn on the bank, the amount being transferred by the bank

[16] See, e.g., Rolnick and Weber (1983).

> to the creditor's account. If the creditor was not a client of the bank, the debtor furnished him with a *fede di credito*, by means of which the latter could pay a third or possibly a fourth or fifth person, who had to be a costumer of the bank. These *fedi di credito* began to circulate. Since [the Italian public banks] regularly permitted credit to be opened on current account (and in Naples mortgage or other loans too), the banks actually created money in the form of paper of several kinds."

In other cases, even though they did not issue notes, banks played a key role in the circulation of IOUs (promises of payment) by acting as clearing houses for these bills (Van der Wee, 1977).

It is worth observing that our results would also obtain in an environment where tangible payment instruments do not necessarily circulate. Indeed, what is important for our hypothesis is that tangible objects are essential in transactions and banks are essential to render these objects valuable. For instance, we could work with an environment where payments are made through cheques and, in turn, the acceptability of cheques relies on the ability of the payee to observe the payer's funds that back the cheque. The presence of a visible institution (bank) where the payer's funds are deposited would allow this. The use of cheques in Early Modern Europe is well documented by Van der Wee (1977).

# 6 Conclusion

In this paper, we have put forward a theory for the origin of banks based on the role of information in trade. The paper studies a decentralized economy where information on transactions is scarce and trade expands over time. We have shown that in such an economy trade can initially be sustained by "fairs", institutions where transactions are observed and recorded publicly. We have then demonstrated that the expansion of trade can progressively undermine the ability of fairs to disseminate information about non-cooperative behavior. When trade intensifies, institutions can continue to sustain it only if they start issuing notes and acting as banks. The sustainability of bank-based exchange hinges on the same feature of institutions that initially supported fair-based exchange, that is their visibility, but it is not undermined by trade expansion. We have argued that the model broadly matches the decline of fairs and the emergence of banks in Medieval and Early Modern Europe. We have also argued that it can offer insights into the impact that the information revolution is having on the role of banks in the payments system.

# References


[1] Aliprantis, C. D., G. Camera, and D. Puzzello (2007) "Anonymous Markets and Monetary Trading," Journal of Monetary Economics 54, 1905-1928.

[2] Araujo, L. and B. Camargo (2006) "Information, Learning, and the Stability of Fiat Money," Journal of Monetary Economics 53, 1571-1591.

[3] Botticini, M. and Z. Eckstein (2005) "Jewish Occupational Selection: Education, Restrictions, or Minorities?," Journal of Economic History 65, 922-48.

[4] Cavalcanti, R., A. Erosa and T. Temzelides (1999) "Private Money and Reserve Management in a Random-Matching Model," Journal of Political Economy 107, 929-945.

[5] Cavalcanti, R. and N. Wallace (1999) "A Model of Private Bank-Note Issue," Review of Economic Dynamics 2, 104-136.

[6] De Roover, R. (1948) Money, Banking and Credit in Mediaeval Bruges, The Mediaeval Academy of America, Cambridge, Massachusetts.

[7] De Roover, R. (1965) "The Organization of Trade," in Postan, M. M., E. E. Rich and E. Miller, eds.: The Cambridge Economic History of Europe, 42-118.

[8] Diamond, D. (1984) "Financial Intermediation and Delegated Monitoring," Review of Economic Studies 51, 393-414.

[9] Ellison, G. (1994) "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching," Review of Economic Studies 61, 567-588.

[10] Friedman, M. (1959) A Program for Monetary Stability, Fordham University Press, New York.

[11] Hayek, F. (1990) Denationalisation of Money - The Argument Refined, The Institute of Economic Affairs, London.

[12] Kandori, M. (1992) "Social Norms and Community Enforcement," Review of Economic Studies 59, 63-80.

[13] Klein, B. (1974) "The Competitive Supply of Money," Journal of Money, Credit, and Banking 6, 423-453.

[14] Kocherlakota, N. (1998) "Money is Memory," Journal of Economic Theory 81, 232-51.

[15] Kreps, D. M. and R. Wilson (1982) "Sequential Equilibria," Econometrica 50, 863-894.

[16] Milgrom, P., D. North, and B. Weingast (1990) "The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and Champagne Fairs," Economics and Politics 2, 1-23.

[17] Ritter, J. (1995) "The Transition from Barter to Fiat-Money," American Economic Review 85, 134-149.

[18] Rolnick, A. and W. Weber (1983) "New Evidence on the Free Banking Era," American Economic Review 73, 1080-1091.

[19] Taub, B. (1985) "Private Fiat Money with Many Suppliers," Journal of Monetary Economics 16, 195-208.

[20] Verlinden, C. (1965) "Markets and Fairs," in Postan, M. M., E. E. Rich and E. Miller, eds.: The Cambridge Economic History of Europe, 199-153.

[21] Williamson, S. (1999) "Private Money," Journal of Money, Credit, and Banking 31, 469-491.

# 7 Appendix

**Proof of Proposition 1.** We start by establishing conditions under which point (ii) of the norm holds. We first show that, for any $\kappa > 0$, there exists a unique $\beta(\kappa)$ such that (for $i \in \{0, ..., n-1\}$)

$$\frac{\beta\left[\frac{1}{2}n(u-c)-a\right]}{1-\beta} \geq (nu+\beta\kappa c-a)\sum_{s=0}^{\infty}\beta^{s+1}(1-\kappa)^{\frac{2^{n-i}(2^{sn}+2^{n}-2)}{2^{n}-1}}+\frac{1}{2}(n-i)(u+c)+\beta\kappa c \quad (17)$$

holds if and only if $\beta \geq \beta(\kappa)$. Fix $\kappa \in (0,1]$ and $i \in \{0, ..., n-1\}$. The left-hand side (LHS) of (17) converges to zero when $\beta$ converges to zero and tends to infinity when $\beta$ converges to one. In turn, the right-hand side (RHS) of (17) converges to $\frac{1}{2}(n-i)(u+c)$ when $\beta$ converges to zero and converges to

$$(nu+\kappa c-a)\sum_{s=0}^{\infty}(1-\kappa)^{\frac{2^{n-i}(2^{sn}+2^{n}-2)}{2^{n}-1}}+\frac{1}{2}(n-i)(u+c)+\kappa c<\infty \quad (18)$$

when $\beta$ converges to one. This implies that there exists $\beta_i(\kappa)$ such that (17) holds with the equality sign. Moreover, the first derivative of the RHS of (17) with respect to $\beta$ is

$$\frac{\partial RHS}{\partial \beta}=\left\{\begin{array}{c}\kappa c+(nu-a)\sum_{s=0}^{\infty}(s+1)\beta^{s}(1-\kappa)^{\frac{2^{n-i}(2^{sn}+2^{n}-2)}{2^{n}-1}}+\\ \kappa c\sum_{s=0}^{\infty}(s+2)\beta^{s+1}(1-\kappa)^{\frac{2^{n-i}(2^{sn}+2^{n}-2)}{2^{n}-1}}\end{array}\right\}>0, \quad (19)$$

and the second derivative of the RHS of (17) with respect to $\beta$ is

$$\frac{\partial\left(\frac{\partial RHS}{\partial \beta}\right)}{\partial \beta}=\left\{\begin{array}{c}(nu-a)\sum_{s=0}^{\infty}s(s+1)\beta^{s-1}(1-\kappa)^{\frac{2^{n-i}(2^{sn}+2^{n}-2)}{2^{n}-1}}+\\ \kappa c\sum_{s=0}^{\infty}(s+1)(s+2)\beta^{s}(1-\kappa)^{\frac{2^{n-i}(2^{sn}+2^{n}-2)}{2^{n}-1}}\end{array}\right\}>0. \quad (20)$$

This implies that the RHS of (17) is strictly convex in $\beta$. Since the LHS of (17) is also strictly convex in $\beta$, there exists a unique $\beta_i(\kappa)$ that satisfies (17) with the equality sign and such that the LHS of (17) weakly exceeds the RHS if and only if $\beta \geq \beta_i(\kappa)$. Henceforth, we let $\beta(\kappa) = \max_{i\in\{0,...,n-1\}} \beta_i(\kappa)$. Next, we demonstrate that, for any $\beta$, if $\kappa$ is sufficiently small the right-hand side of (17) exceeds the left-hand side and agents have no incentive to cooperate, for all $i \in \{0, ..., n-1\}$. To show this, we simply need to evaluate the RHS of (17) at $\kappa = 0$. Now, because

$$\frac{\beta\left[\frac{1}{2}n(u-c)-a\right]}{1-\beta}<\frac{\beta(nu-a)}{1-\beta}, \quad (21)$$

condition (17) is violated at $\kappa = 0$. By continuity, it must be that this condition is also violated for $\kappa$ sufficiently close to zero.

We now prove that (17) implies that an agent wants to cooperate inside institutions. Consider an agent in a meeting with another agent inside an institution in period $t$. If she cooperates, her expected flow payoff is $\frac{1}{2}(1-\beta)(1+n)(u-c)$ while her expected continuation payoff is $v_f$. If she deviates, she obtains $(1-\beta)u$ and moves to autarky (because under the social norm, as soon as a deviation occurs inside an institution, all agents move to autarky). In all future periods, the agent cooperates inside institutions and stays in autarky, which implies a continuation payoff of $\frac{3}{2}\kappa(u-c)+a$. Therefore, the agent will follow the social norm if and only if

$$\frac{1}{2}n(u-c)\geq a+(1-\beta)\frac{1}{2}(u+c). \quad (22)$$

Applying a similar reasoning to an agent in a meeting with the manager of an institution, we obtain that the agent will follow the social norm and cooperate if and only if

$$\frac{1}{2}n\left(u-c\right)\geq a+\left(1-\beta\right)c. \tag{23}$$

Note that condition (17) can be rewritten

$$\frac{1}{2}n\left(u-c\right)\geq a+\frac{1}{\beta}\left[\begin{array}{c}\left(1-\beta\right)\left(nu+\beta\kappa c-a\right)\sum_{s=0}^{\infty}\beta^{s+1}\left(1-\kappa\right)^{\frac{2^{n-i}\left(2^{sn}+2^{n}-2\right)}{2^{n}-1}}\\+\frac{1}{2}\left(n-i\right)\left(u+c\right)+\beta\kappa c\end{array}\right], \tag{24}$$

which certainly implies (22) and (23).

We next turn to establish conditions for iii. If she has observed a deviation in a private meeting but no deviation inside institutions, an agent participates in private exchange but does not cooperate in private meetings or inside institutions. We start by proving that an agent does not cooperate inside institutions. The strongest incentive to slow down the spread of non-cooperative behavior occurs when (1) the agent meets the manager of an institution after observing a deviation for the first time in her last private meeting in the previous period; and (2) she believes that this is the first deviation that occurs in the economy. Consider then the decision problem of an agent in this scenario. If she follows the social norm, in each period she deviates with the manager (obtaining $u$), cooperates with another agent inside the institution (obtaining in expectation $\frac{1}{2}\left(u-c\right)$) and then stays in autarky (because after her deviation with the manager all agents move to autarky) obtaining $a$. Hence, her expected payoff is $\left(1-\beta\right)\left[u+\frac{1}{2}\left(u-c\right)+a\right]+\beta\left[\frac{3}{2}\kappa\left(u-c\right)+a\right]$. If, instead, she decides to slow down the spread of non-cooperative behavior, there is a probability $\kappa$ that a deviation will occur inside an institution (i.e. the probability that the agent she met at the end of the previous period enters an institution), in which case her expected payoff is $\left(1-\beta\right)\frac{3}{2}\left(u-c\right)+\beta\frac{3}{2}\kappa\left(u-c\right)+a$. There is a complementary probability that such a deviation will not occur and her expected payoff is $\left(1-\beta\right)\frac{3}{2}\left(u-c\right)+v_2$. Comparing her expected payoffs, the agent will follow the social norm and deviate if and only if

$$\left(1-\beta\right)c\geq\left(1-\kappa\right)\left(v_2-v_d\right), \tag{25}$$

which, using the expressions for $v_2$ and $v_d$ in the main text, can be rewritten as

$$\sum_{s=0}^{\infty}\beta^{s}\left(1-\kappa\right)^{\frac{2\left(2^{sn}-1\right)+2^{n}-1}{2^{n}-1}}\leq\frac{c}{nu+\beta\kappa c-a}. \tag{26}$$

Fix $\beta\in(0,1)$. The left-hand side (LHS) of (26) converges to $\frac{1}{1-\beta}$ when $\kappa$ converges to zero and converges to zero when $\kappa$ converges to one. Moreover, the LHS of (26) is strictly decreasing in $\kappa$. In turn, the right-hand side (RHS) of (26) is strictly decreasing in $\kappa$, it converges to $\frac{c}{nu-a}$ when $\kappa$ converges to zero while it is equal to $\frac{c}{nu+\beta c-a}$ when $\kappa$ is equal to one. Because $\frac{c}{nu-a}<1\leq\frac{1}{1-\beta}$, there exists a unique $\kappa^{*}\left(\beta\right)$ such that (26) holds for all $\kappa\geq\kappa^{*}\left(\beta\right)$. Turning next to the claim that an agent has no incentive to cooperate in private meetings, this certainly holds under (26). In fact, if an agent is not tempted to slow down non-cooperative behavior inside an institution *a fortiori* she will not be tempted to slow it down in private meetings.

Finally, we also need to establish conditions for iv (a) If she has observed deviations inside institutions once, an agent stays in autarky but cooperates inside institutions; (b) if she has observed deviations inside institutions more than once, an agent stays in autarky and she does not cooperate inside institutions. Assume that a deviation inside institutions has already occurred in period $t$ and consider the incentives to follows the social norm inside an institution in some period $t'>t$. First, if the deviation in period $t$ is the only deviation that has ever happened in the economy, the social norm implies that there is no cooperation outside institutions but there is cooperation inside them. In this case, if an agent cooperates with a manager, she expects a payoff $(1-\beta)\frac{3}{2}(u-c)+\beta\frac{3}{2}\kappa(u-c)+a$. If she deviates, she obtains $(1-\beta)u+\beta(1-\beta)\kappa u+a$ because there will be no cooperation by agents in any future meeting. The agent will not deviate if and only if

$$\beta \geq \frac{3c-u}{3c-u+\kappa(u+2\beta u-3c)}. \tag{27}$$

An agent may also have an incentive to deviate in a meeting with another agent. If she cooperates, she expects $(1-\beta)\frac{1}{2}(u-c)+\beta\frac{3}{2}\kappa(u-c)+a$, while a deviation implies an expected payoff of $(1-\beta)u+\beta(1-\beta)\kappa u+a$. The agent will cooperate if and only if

$$\beta \geq \frac{u+c}{u+c+\kappa(u+2\beta u-3c)}. \tag{28}$$

Clearly, (28) implies (27). Hence, for any given $\kappa$, there exists a unique $\beta'(\kappa)$ such that (28) holds if and only if $\beta \geq \beta'(\kappa)$. Finally, point iii (b) holds because an agent has no incentive to cooperate since cooperation would reduce her flow payoff and would not change her continuation payoff. Summarizing, for any given $\beta$ and $\kappa$, there exists $\beta^*(\kappa) \equiv \max\left\{\beta(\kappa), \beta'(\kappa)\right\}$ and $\kappa^*(\beta)$ such that the social norm is an equilibrium if and only if $\beta \geq \beta^*(\kappa)$ and $\kappa \geq \kappa^*(\beta)$.

It remains to be shown that there exists a non-empty region of parameters such that $\beta \geq \beta^*(\kappa)$ and $\kappa \geq \kappa^*(\beta)$. Note that the LHS of (26) is increasing in $\beta$ and the RHS of (26) is decreasing in $\beta$. Moreover, when $\beta$ converges to 1, we have

$$\sum_{s=0}^{\infty}(1-\kappa)^{\frac{2(2^{sn}-1)+2^n-1}{2^n-1}} \leq \frac{c}{nu+\kappa c-a}. \tag{29}$$

Let $\overline{\kappa}$ be such that

$$\sum_{s=0}^{\infty}(1-\overline{\kappa})^{\frac{2(2^{sn}-1)+2^n-1}{2^n-1}} = \frac{c}{nu+\overline{\kappa} c-a}. \tag{30}$$

This implies that, for all $(\beta,\kappa)$ such that $\kappa \geq \overline{\kappa}$ and $\beta \geq \beta^*(\kappa)$, the social norm is an equilibrium.

**Proof of Proposition 2.**

Existence of steady state equilibrium.

We have to prove that there exists a steady state equilibrium in which agents always participate in private exchange, $\mathsf{V}=(1,0,1,0,1,0)$, i.e.: (i) agents always produce in exchange for notes ($\lambda_1=1$) but never produce if they are not paid with notes ($\lambda_0=0$), (ii) inside institutions notes are always redeemed ($\gamma_1=1$) while agents without notes do not receive any good from a manager ($\gamma_0=0$), (iii) managers offer notes in exchange for goods to all agents who have just redeemed a note, and only to them ($\alpha_1=1$ and $\alpha_0=0$). Moreover, if in any period a manager deviates from the behavior described by the vector $\mathsf{V}=(1,0,1,0,1,0)$ no agent produces from that period on. Before proving our claim, it is useful to display again agents' value functions implied by the candidate steady-state $\mathsf{V}=(1,0,1,0,1,0)$ (for $i\in\{0,...,n\}$, where $i=0$ corresponds to the meeting inside an institution):

$$v_1^i(m)=mv_1^{i+1}(m)+(1-m)\left[u+v_0^{i+1}(m)\right], \tag{31}$$

$$v_0^i(m)=m\left[-c+v_1^{i+1}(m)\right]+(1-m)v_0^{i+1}(m), \tag{32}$$

$$v_0^{n+1}(m)=\beta\left[\kappa v_0^0(m)+(1-\kappa)v_0^1(m)\right], \tag{33}$$

$$v_1^{n+1}(m)=\beta\left\{\kappa\left[u-c+v_1^0(m)\right]+(1-\kappa)v_1^1(m)\right\}. \tag{34}$$

First, an agent is always willing to produce in exchange for a note as long as $-c + v_1^i(m) \geq v_0^i(m)$, for all $i \in \{0, ..., n\}$, and $-c + v_1^{n+1}(m) \geq v_0^{n+1}(m)$. Using (31)-(34), the first inequality always holds while the second holds as long as

$$\beta \geq \frac{c}{c + (1-m)(u-c) + \kappa(u-c)}. \tag{35}$$

We also need to check under what conditions an agent prefers engaging in private exchange than staying in autarky. Clearly, $-c + v_1^0(m) \geq v_0^0(m)$ implies that it is always better to engage in private exchange with a note than without a note. Hence, we only need to check the incentives of an agent without a note. From (31) and (32), the expected payoff of engaging in private exchange without a note is

$$v_0^1(m) = v_0^n(m) + (n-1)m(1-m)(u-c), \tag{36}$$

where

$$v_0^n(m) = v_0^{n+1}(m) - mc + m\beta[\kappa(u-c) + (1-m)u + mc]. \tag{37}$$

An agent without a note prefers engaging in private exchange than staying in autarky if and only if $v_0^1(m) \geq a + v_0^{n+1}(m)$, that is

$$-mc + (n-1+\beta)m(1-m)(u-c) + \beta m[\kappa(u-c) + c] \geq a. \tag{38}$$

It remains to be shown that each manager is always willing to behave according to $\mathsf{V} = (1, 0, 1, 0, 1, 0)$. This is immediate because all agents stop producing as soon as they observe that a manager is issuing more notes than the prescribed amount $m$. Summarizing, there exists a steady state equilibrium in which agents always participate in private exchange, the measure of agents with notes equals $m$ and agents and managers behave according to $\mathsf{V} = (1, 0, 1, 0, 1, 0)$ if and only if

$$\beta \geq \frac{c}{c + (1-m)(u-c) + \kappa(u-c)} \tag{39}$$

and

$$-mc + (n-1+\beta)m(1-m)(u-c) + \beta m[\kappa(u-c) + c] \geq a. \tag{40}$$

The conditions in Proposition 2 imply (39) and (40).

Transition to steady state.

We demonstrate here that there exists a transition from an initial state in which no agent holds a note to a steady state in which a measure $m$ of agents hold a note. The argument runs as follows. Consider the following behavior of agents and managers. In all periods up to period $t-1$, all agents who meet a manager without a note receive a note; all agents who meet a manager without a note at the beginning of period $t$ receive a note with probability $q$. Finally, in all periods up to period $t$ there is no redemption of notes and agents do not need to produce in exchange for a note. If agents and managers behave this way, at the beginning of period $t$ the measure of agents with notes is

$$m_t = \kappa + (1-\kappa)\kappa + ... + (1-\kappa)^{t-1} q\kappa, \tag{41}$$

which can be rewritten as

$$m_t = 1 - (1-\kappa)^{t-1}(1-q\kappa). \tag{42}$$

For all $q \in [0, 1]$, $\frac{\partial m_t}{\partial t} > 0$ and $\lim_{t\to\infty} m_t = 1$. Moreover, if $t = 1$ and $q = 0$, $m_1 = 0$. Hence, for every $m \in [0, 1)$, there exists $q_m \in [0, 1]$ and $t_m \in \mathbb{N}$ such that

$$m = 1 - (1-\kappa)^{t_m - 1}(1-q_m\kappa). \tag{43}$$

We can then define a note issuing path $\{m_t\}_{t=1}^{\infty}$ given by

$$m_t = \begin{cases} 1-(1-\kappa)^t \text{, if } t < t_m \\ 1-(1-\kappa)^{t_m-1}(1-q_m\kappa) = m \text{, if } t = t_m \\ m \text{, if } t > t_m \end{cases} . \tag{44}$$

This path converges to $m$ in a finite number of periods $t_m$. Now, consider the economy in period $t_m+1$, when the measure of agents with notes equals $m$. Consider the following behavior of agents in $t_m+1$. If an agent with a note meets a manager, she produces for the manager. This implies that, starting in period $t_m+2$, the measure of notes that are presented for redemption equal the measure of goods stored by managers. From period $t_m+2$ onwards, the economy is in a steady-state in which, as long as agents and managers behave according to $\mathsf{V}=(1,0,1,0,1,0)$ and as long as (35) and (38) hold, notes are valued as a medium of exchange and the amount of notes in circulation equals $m$. It remains to be shown that the transition just described arises as an equilibrium outcome. Because the transition is implemented through a sequence of specified actions that occur inside institutions from period 1 to period $t_m+1$, all agents in the economy observe a deviation from a specified action. The following punishment scheme can then be implemented: if an agent or a manager deviates from the sequence of specified actions, no agent will produce again, inside or outside institutions. Otherwise, agents and managers continue to behave consistently with the specified actions and, after the transition is complete, agents and managers behave according to $\mathsf{V}=(1,0,1,0,1,0)$. Clearly, because the continuation payoff if there is a deviation equals zero and the continuation payoff if there is no deviation is positive, no agent or manager has an incentive to deviate from the actions specified in the transition path.

**Proof of Proposition 4.** The argument in this proof is developed for $\theta=\varepsilon=0$. By continuity, it also holds for $\theta$ and $\varepsilon$ sufficiently close to zero. Consider a strategy profile under which managers behave as in point i.(adapted) and agents behave as in points (iii) and (iv)(adapted) while they behave as in point (ii) if and only if condition (16) is satisfied. This strategy profile is an equilibrium as long as each agent believes that all other agents will always behave as in point (ii) and cooperate on the equilibrium path. This belief can be justified as follows. Consider first an agent with production cost $c$. If $\theta=0$, she believes that all other agents have her same production cost. A similar reasoning applies to an agent with cost $c_H$. If $\varepsilon=0$, she believes that the state $\mathcal{S}_1$ is realized, in which case she believes that she is the only agent with a high production cost. In sum, when $\theta=\varepsilon=0$, regardless of her production cost, an agent believes that all other agents have cost $c$. This implies that each agent believes that all other agents always behave as in point (ii) as long as condition (16) holds for $\widetilde{c}=c$. Now, let us show that indeed we can always find values $\overline{t}$ and $t(\beta,c_H)\le\overline{t}$ such that $c_H$ agents (and, hence, $c$ agents) follow (iv)(adapted), that is

$$\beta \geq \frac{c_H}{c_H+(1-m)(u-c_H)}, \tag{45}$$

$$nm(1-m)(u-c_H) \geq a+(1-\beta)m[(1-m)u+mc_H]; \tag{46}$$

$c_H$ agents behave as in point (ii) up to $t(\beta,c_H)$ and deviate at $t(\beta,c_H)$, that is, for an arbitrarily small $\epsilon>0$,

$$\frac{\beta}{1-\beta}\begin{bmatrix} \widetilde{v}_{f,t+1} \\ -\widetilde{w}_{t+1}(m) \end{bmatrix} \binom{>}{=} \begin{Bmatrix} \displaystyle\sum_{s=0}^{\infty}\beta^{s+1}\prod_{j=1}^{s+1}\left(1-\frac{\kappa}{\eta_{t+j}}\right)^{2^{jn-i}}\begin{bmatrix} nu-\widetilde{w}_{t+s+2}(m) \\ +\beta\frac{\kappa}{\eta_{t+s+2}}u \end{bmatrix} \\ +\frac{1}{2}(n-i)(u+c_H)+\frac{\beta\kappa}{2\eta_{t+1}}(3u-c_H)-\epsilon \end{Bmatrix} if\ t\binom{<}{=}t(\beta,c_H), \tag{47}$$

while $c$ agents always behave as in point (ii), that is

$$\frac{\beta}{1-\beta}\left[\widetilde{v}_{f,t+1}-\widetilde{w}_{t+1}\left(m\right)\right] > \left\{\begin{array}{c} \sum_{s=0}^{\infty}\beta^{s+1}\prod_{j=1}^{s+1}\left(1-\frac{\kappa}{\eta_{t+j}}\right)^{2^{jn-i}}\left[nu-\widetilde{w}_{t+s+2}\left(m\right)+\beta\frac{\kappa}{\eta_{t+s+2}}u\right] \\ +\frac{1}{2}\left(n-i\right)\left(u+c\right)+\frac{\beta\kappa}{2\eta_{t+1}}\left(3u-c\right) \end{array}\right\} \forall t; \tag{48}$$

and all agents behave as in point (iii), that is (for $\widetilde{c}\in\{c,c_H\}$)

$$\left(1-\frac{\kappa}{1+\eta(t-1)}\right)\sum_{s=0}^{\infty}\prod_{j=2}^{s+1}\left(1-\frac{\kappa}{\eta_{t+1+j}}\right)^{2^{jn-n+1}}\left[nu-\widetilde{w}_{t+s+2}\left(m\right)+\frac{\kappa}{\eta_{t+s+3}}u\right]\leq\widetilde{c}\;\;\forall t. \tag{49}$$

First of all, observe that we can always set $\eta$ low enough and $\kappa$ high enough that condition (49)) holds. Second, for simplicity, let $a$ equal zero. Then, condition (45) implies condition (46). Next, let condition (45) hold with the equality sign, that is $\beta=\frac{c_H}{c_H+(1-m)(u-c_H)}$. By setting $m$ and $c_H$ sufficiently low, we can render $\beta$ and, hence, the left hand side of (47) arbitrarily small while the right hand side of (47) can be made sufficiently larger than zero by appropriately choosing $n$ and $u$. It is also immediate that by letting condition (45) be slack and setting $\beta$ sufficiently high, the left hand side of (47) becomes arbitrarily large while the right hand side stays finite. By continuity of both sides of (47), all this guarantees that for any chosen $t$ we can find parameter values such that condition (47) holds with the equality sign. Moreover, since for $t=1$ the value of $\beta$ that satisfies condition (47) with the equality sign is strictly below one, we can find values $t\left(\beta,c_H\right)$ such that condition (47) holds with the inequality sign for $t<t\left(\beta,c_H\right)$ and with the equality sign for $t=t\left(\beta,c_H\right)$. Finally, condition (48) is always looser than (47) for $\epsilon$ sufficiently small. Therefore, (48) holds at $t\left(\beta,c_H\right)$ and, in addition, one can always select $\overline{t}\geq t\left(\beta,c_H\right)$ such that it continues to hold thereafter.