Information, Trade and the Origin of Banks

Luis Araujo  
Michigan State University

Raoul Minetti  
Michigan State University

This Draft: April 2008

Abstract

This paper puts forward an explanation for the origin of banks based on the role of information in trade. We study an economy in which agents trade inside institutions, where transactions are publicly observable, and in private meetings. As long as trade is limited, institutions can act as "fairs" and sustain it by disseminating information on agents’ (non-)cooperative behavior. However, as fairs boost trade the speed at which they disseminate information becomes too low relative to the speed of trade. Institutions can then continue to sustain trade only by acting as banks and issuing notes for the payment of transactions. The visibility of institutions keeps on being crucial because it prevents banks from overissuing notes. The model broadly matches historical evidence on the origin of banks in Medieval and Renaissance Europe.

Keywords: Banking, Information, Trade.

1 Introduction

The origin of banks constitutes a fascinating and yet relatively unexplored issue. This paper puts forward an hypothesis for this origin based on the role of information in trade. We explain banks, meant as institutions that sustain trade by issuing a medium of exchange (“notes”), as the evolution of institutions that sustain trade without the need of media of exchange. Our hypothesis can be summarized as follows. We study an economy where agents trade inside and outside institutions. The feature that characterizes institutions is 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 in the economy, as soon as such a behavior reaches an institution it gets publicly disclosed. In this first stage, institutions thus act as “fairs”, locations where transactions occur and are recorded publicly. The ability of fairs to disseminate information about the lack of cooperation in the economy hinges however on the physical occurrence and recording of transactions inside them. As trade expands, the ability of institutions to host and record transactions gets strained. Therefore, institutions can continue to sustain trade only if they start acting as “banks” and issuing an object (notes) that immediately delivers payment of transactions. The public observability
of transactions occurring inside institutions continues to be crucial at this stage: it is exactly because transactions inside them are publicly observable that institutions can commit not to overissue notes. Crucially, unlike in the case of fairs, the role of banks is not undermined by the expansion of trade 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 - the issue of notes containing all the information required for transactions - does not suffer from this problem because notes are an object that can be issued at no cost.

To gain a better understanding of the hypothesis it is useful to delve deeper into our model. When institutions act as fairs, trade operates through a norm of gift-exchange. Under this norm, an agent produces for another agent outside a fair only if she has always observed such a cooperative behavior in the past. This norm entails a contagion mechanism: if an agent does not cooperate, she will initiate a sequence of non-cooperative behaviors throughout the economy which may eventually backfire on herself. Absent fairs, agents would have no incentive to follow the norm. In fact, contagion would be too slow and an agent who does not cooperate would never be reached by the contagion process and bear the consequences of her behavior. In contrast, fairs render the norm sustainable because they dramatically accelerate the contagion process: as soon as the non-cooperative behavior reaches a fair it gets disclosed to the whole economy and backfires on the initial defector.

By providing proper incentives for cooperation, fairs allow new agents to be progressively integrated into the exchange process. Yet, as anticipated above, trade expansion contains the germs for their demise and the transition to banks. As trade expands, the time it takes for a non-cooperative behavior to reach a fair lengthens. This implies that an agent will bear the consequences of her non-cooperative behavior far in the future, which dilutes her incentive to cooperate. At this stage, the circulation of a tangible object (notes) that immediately delivers payment becomes essential to sustain trade. However, since transactions outside institutions are not observable, agents have an incentive to overissue notes, driving their value to zero. This problem can only be overcome by institutions because transactions occurring inside them - and hence their overissue of notes - are publicly observable. In sum, the same observability of transactions that previously allowed institutions to act as fairs (disseminating information about agents’ lack of cooperation) now allows them to act as banks (issuing notes). It is in this sense that banks constitute the natural evolution of fairs.

The remainder of the paper is organized into four sections. In Section 2, we relate the paper to the literature. Section 3 presents the model. In Section 4, we discuss some historical facts from Medieval and Renaissance Europe that are consistent with our hypothesis. Section 5 concludes. Formal details of the proofs are relegated to the Appendix.
2 Prior Literature

This paper contributes to the literature on the role and origin of banks. While it is widely agreed that banking arose from the evolution of mercantile activity, the explanations of this evolution differ. Some scholars stress that some merchants progressively accumulated capital in trades and at some point this enabled them to start offering credit to other merchants. Other scholars focus on the accumulation of human capital rather than financial capital: by engaging in trades, merchants acquired skills and knowledge that eventually allowed them to perform the sophisticated activities of bankers (for an extensive discussion of both these hypotheses see, e.g., De Roover, 1948). These explanations radically differ from ours.

Our paper also relates to the literature on the sustainability of cooperation in trades. Although this literature does not focus on the origin of banks, we can carry out 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 is sustainable because the population is finite and a deviation triggers a contagion process that eventually backfires on the defector. In Kandori (1992) and Ellison (1994) institutions play no role. In our environment, instead, the population is infinite and institutions are crucial to spread information and accelerate the contagion process. In Milgrom, North and Weingast (1990), fairs sustain trade because they act as “primitive courts”. In fact, they are able to collect information about (monitor) behavior in transactions among merchants and punish non-cooperative merchants. Although Milgrom, North and Weingast (1990) are silent about banks, it is not difficult to notice the similarities of their approach with the theory of banks as monitors of Diamond (1984). In Diamond (1984), banks arise because they have a superior ability to monitor and punish the misbehavior of borrowers. One could then view banks à la Diamond (1984) as the evolution of fairs à la Milgrom, North and Weingast (1990).

The theory of fairs we propose differs from that in Milgrom, North and Weingast (1990). In our decentralized environment information on transactions among agents is private and no institution can act as a monitor. The implications for the origin of banks that stem from our hypothesis are also different. In our environment, fairs evolve towards institutions that issue notes. Our emphasis is then on the origin of banks as “mechanisms of settlement” rather than as “credit providers or monitors”. In this sense, our interpretation of banks is close to that in Cavalcanti and Wallace (1999) in which bank notes circulate as a medium of exchange because banks are visible. In our context, this visibility is key to prevent overissue of notes.2

---

1 Botticini and Eckstein (2005) focus on the involvement of Jews in mercantile and banking activities and stress the role of literacy among Jews in fostering the accumulation of skills necessary for these activities.

2 Dating back at least to Milton Friedman (1959), it is well known that, in the absence of a control on the amount of notes in circulation, a probable 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 note issue for reputational concerns (see also Araujo and Camargo (2006)); Taub (1982) and Hayek (1990) argue...
3 The Model

In this section, we first describe the environment. Then, we analyze how trade can be sustained by a norm of gift-exchange when institutions act as fairs. Next, we analyze how trade can be sustained by the circulation of tangible objects (notes) when institutions act as banks. Finally, we put forward an endogenous transition from fairs to banks driven by the expansion of trade.

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 a manager. There are perishable and indivisible consumption goods. Each agent can produce one good at a cost of $c$ per unit. An agent does 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. A manager cannot produce but she can store up to a measure one of goods across periods, deriving 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. Both agents and managers discount across periods at a rate $\beta \in (0, 1)$.

At the beginning of every period, each institution hosts a measure $\mu \in (0, 1]$ of randomly drawn agents. Inside the institution, such agents first meet the manager and then randomly and bilaterally meet with each other. After agents meet 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 (produce), a random draw defines who is the consumer (producer).

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, they are only observable by the agents who engage in the meeting.

3.2 Institutions as Fairs

In this section, we show that institutions can sustain trade even if no agent or manager issues notes. Institutions can do so because, thanks to their visibility, they disseminate information about agents' that competition among various suppliers of notes discipline note issuers; Ritter (1995) shows that a coalition of note issuers do not overissue because they internalize the negative effects of overissue in terms of a deterioration of the value of the notes. Cavalcanti and Wallace (1999), Williamson (1999), Berentsen (2006), and Martin and Schreft (2006) show that overissue and/or a refusal to redeem notes can be avoided when note issuers are monitored.
(non-)cooperative behavior. We simplify the description of actions by assuming that, upon entering a meeting with another agent, each agent announces “produce” or “consume”. This announcement is made before the agent knows whether she will be a producer or not and it is binding within the meeting. We say that an agent cooperates if and only if she announces “produce”. Accordingly, in the meetings between agents and the manager of an institution, we say that an agent cooperates if she produces for the manager and that the manager cooperates if she redistributes her stored goods among the agents.

In what follows, we prove that an equilibrium exists in which managers always cooperate and all agents behave in accordance with the following norm of gift-exchange:

An agent cooperates at the beginning of period 1; 

i. (Cooperation) If cooperation is the only outcome she has ever observed, an agent participates in private exchange and cooperates both inside institutions and in private meetings;

ii. (Deviations in private meetings) If she has observed a deviation in a private meeting but no deviation inside institutions: (a) an agent participates in private exchange but does not cooperate in private meetings, (b) an agent deviates in her first meeting with the manager of an institution;

iii. (Deviations inside institutions) (a) If she has observed deviations inside institutions once, an agent stays in autarky but she continues to cooperate inside institutions, (b) if she has observed deviations inside institutions more than once, an agent stays in autarky and does not cooperate with defectors inside institutions. Finally, any agent who cooperates with a defector becomes herself a defector.

The norm implies that cooperation in private meetings is sustained by a form of community enforcement: a deviation in a private meeting triggers a chain of deviations that, upon reaching an institution, collapses all exchange in private meetings, backfiring on the initial defector (see point (ii)). Unlike in Kandori (1992), where the population is finite, in our environment such a form of community enforcement hinges on the existence of institutions where transactions are publicly observable. Note that, since in our environment transactions in private meetings cannot be publicly observed, personal retaliation cannot sustain cooperation in private meetings. In contrast, personal retaliation can sustain cooperation inside institutions. In fact, agents can keep a record of the defectors inside institutions and punish them (see point (iii)).

If all agents cooperate, at the beginning of period $t > 1$ an agent’s (normalized) expected payoff from following the norm of gift-exchange is

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

(1)

$^3$Since an agent can announce “consume”, this assumption does not violate the agent’s participation constraint. As it will become clear shortly, this assumption simplifies the computation of payoffs after a deviation from the equilibrium path.
which can be rewritten as $[k\mu + \frac{1}{2}(k\mu + n)](u - c)$. In each period, there is a probability $k\mu$ that 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 $(a)$. We prove the existence of an equilibrium in which the norm of gift-exchange holds in steps, starting with the incentives of agents to follow the norm on the equilibrium path. Henceforth, we let $\kappa \equiv k\mu$.

i. If cooperation is the only outcome she has ever observed, an agent participates in private exchange and cooperates both inside institutions and in private meetings. We start investigating the conditions under which an agent has the incentive to cooperate in private meetings. 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 norm and cooperates, her expected flow payoff is $\frac{1}{2} (1 - \beta) (n - i) (u - c)$ and her expected continuation payoff is $v_f$. If she deviates, her expected 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 enter an institution at the beginning of period $t + 1$, in which case the agent’s expected continuation payoff is $\frac{1}{2} (1 - \beta) (n - i) (u - c)$. 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, her flow payoff inside the institution is $u + \frac{1}{2} (u - c)$ because under the norm, the agent does not cooperate with the manager (obtaining $u$), cooperates with another agent (obtaining in expectation $\frac{1}{2} (u - c)$), and during the period stays in autarky (obtaining $a$). Moreover, using a similar reasoning, the agent’s continuation payoff is $\beta [\frac{1}{2} \kappa (u - c) + a]$. There is a complementary probability that none of the $2^{n - i}$ agents reached by the contagion process will enter an institution in period $t + 1$ and the expected 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 not deviate if and only if (for $i \in \{0, ..., n - 1\}$)

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

(2)

where $v_d \equiv a + \beta \frac{1}{2} \kappa (u - c)$ is the expected payoff at the beginning of private meetings when a deviation inside institutions was observed once (see (ii)). The computation of $v_{2^{n - i}} - v_d$ 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 expected flow payoff from following the norm is $-k\mu c + \frac{1}{2} (k\mu + n) (u - c)$. At the

\footnote{At the beginning of period 1, managers have no goods in storage and the (normalized) expected payoff from following the norm is $-k\mu c + \frac{1}{2} (k\mu + n) (u - c)$.}
participates in exchange but she does not cooperate in private meetings. If she has observed a deviation in a private meeting but no deviation inside institutions, meeting with the manager of an institution, we also prove that \((5)\) implies that an agent prefers engaging in private exchange than staying in autarky. This completes the demonstration of point (i).

Moreover, there exists a unique \(\kappa\) such that condition (5) holds if and only if \(\beta \geq \beta (\kappa)\). 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. 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. It is also easy to see that for any given \(\beta\), 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 is small and the contagion process will be slow. Finally, observe that it is straightforward that the condition under which an agent has the incentive to cooperate inside institutions is looser than (5). Intuitively, if an agent deviates inside an institution she will immediately trigger a collapse of private exchange. In the Appendix, we prove this formally and we also prove that (5) implies that an agent prefers engaging in private exchange than staying in autarky. This completes the demonstration of point (i).

We now study the incentive of an agent to follow the norm out of the equilibrium path. We provide here an intuitive discussion while we relegate formal details to the Appendix. ii and iii (a). If she has observed a deviation in a private meeting but no deviation inside institutions: (a) an agent participates in exchange but she does not cooperate in private meetings, (b) an agent deviates in her first meeting with the manager of an institution. Moreover, if she has observed deviations inside institutions once, an agent stays in autarky but she continues to cooperate 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.

\[\frac{\beta \left(\frac{1}{2}n (u - c) - a\right)}{1 - \beta} \geq (nu + \beta \kappa c - a) \sum_{s=0}^{\infty} \beta^{s+1} (1 - \kappa) \frac{2^{n-1} \sum_{k=1}^{n-1} k^n}{2^{n-1} - 1} + A_i (\beta, \kappa), \]

where \(A_i (\beta, \kappa) \equiv \frac{1}{2} (n - i) (u + c) + \beta \kappa c\).
The agent is not tempted to do so if and only if the probability that anyway other agents will spread non-cooperative behavior is sufficiently high. In turn, the probability that other agents deviate inside institutions is higher the larger the number of agents who 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 iii(b), i.e. if she has observed deviations inside institutions more than once, an agent stays in autarky and does not cooperate with defectors inside institutions. The straightforward proof of this point is also presented in the Appendix. We obtain again that neither \( \beta \) nor \( \kappa \) can be too small. Proposition 1 wraps up and summarizes our result.

**Proposition 1** For any given \( \kappa \) and \( \beta \), there exists a discount factor \( \beta^* (\kappa) > 0 \) and a probability that an agent is randomly assigned to meet inside an institution \( \kappa^* (\beta) > 0 \) such that the norm of gift-exchange is an equilibrium if and only if \( \beta \geq \beta^* (\kappa) \) and \( \kappa \geq \kappa^* (\beta) \). In particular, \( \lim_{\beta \to 1} \kappa^* (1) < 1 \) and the norm of gift-exchange is an equilibrium for all \( \beta \geq \beta^* (\kappa) \) and \( \kappa \geq \lim_{\beta \to 1} \kappa^* (1) \).

The proposition has an immediate implication. Consider an environment where the size of the population involved in trade increases over time, in a way such that the probability \( \kappa \) that any given agent enters an institution progressively drops. Eventually, the incentive of an agent to follow the norm of gift-exchange will break down. We will exploit this result later when we study the transition from fairs to banks.

### 3.3 Institutions as Banks

In this section, we demonstrate that institutions can also sustain trade by allowing the circulation of notes as a medium of exchange. In our economy, transactions in private meetings are not 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, all agents will choose autarky at the beginning of every period. 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). Here, we focus on a steady state equilibrium while in the Appendix we also analyze the transition to this steady-state. The behavior of agents and managers is summarized by a vector \( P = (\lambda, \gamma, \eta, \alpha_1, \alpha_0) \), where \( \lambda \) is the probability that an agent produces in exchange for a note, \( \gamma \) is the probability that an outstanding note is redeemed (i.e., transferred from an agent to a manager in exchange for one unit of good), \( \eta \) is the probability that an agent receives goods from a manager if she has no note, \( \alpha_1 \in [0, 1] \) is the probability that an agent who has just redeemed a note receives a new note from a manager, and...
\( \alpha_0 \in [0, 1] \) is the probability that an agent who has not redeemed a note receives a new note from a manager. In the Appendix we prove that there exists a steady state equilibrium with the following features: (i) agents always produce in exchange for notes (\( \lambda = 1 \)), (ii) notes are always redeemed (\( \gamma = 1 \)), (iii) agents without notes do not receive any good from a manager (\( \eta = 0 \)), (iv) managers offer new notes in exchange for a good to all agents who have just redeemed an old note, and only to them (\( \alpha_1 = 1 \) and \( \alpha_0 = 0 \)), (v) agents always participate in private exchange. Moreover, if any manager deviates from \( P \) no agent will produce for this manager in any future meeting and no agent will produce in exchange for notes issued by this manager in any future meeting.

It is useful to display the value functions implied by the candidate steady-state. Let \( m \) denote the measure of agents with notes and \( v^i_j (m) \) 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 inside an institution. Moreover, let \( v^{n+1}_j (m) \) denote the expected payoff of an agent with \( j \) notes at the end of her private meetings in a period. For all \( i \in \{0, ..., n\} \), we obtain

\[
\begin{align*}
    v^i_1 (m) &= mv^{i+1}_1 (m) + (1 - m) \left[ u + v^{i+1}_0 (m) \right], \\
v^i_0 (m) &= m \left[ -c + v^{i+1}_1 (m) \right] + (1 - m) v^{i+1}_0 (m), \\
v^{n+1}_1 (m) &= \beta \left\{ \kappa \left[ u - c + v^0_1 (m) \right] + (1 - \kappa) v^1_1 (m) \right\}, \\
v^{n+1}_0 (m) &= \beta \left[ \kappa v^0_0 (m) + (1 - \kappa) v^1_0 (m) \right].
\end{align*}
\]

Consider \( v^i_1 (m) \) (a similar interpretation holds for \( v^0_0 (m) \)). 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 both agents want to consume but none has the incentive to produce for a note given the upper bound on note holdings. With probability \( 1 - m \), the agent meets instead an agent without a note, transfers her note in exchange for a good, obtaining utility \( u \), and moves to the next meeting without a note. Consider next \( v^{n+1}_1 (m) \). At the end of a period, an agent with a note expects to enter an institution with probability \( \kappa \) at the beginning of the following period, in which case she redeems her note in exchange for one unit of good (obtaining \( u \)), produces for a manager (at the cost \( c \)) and obtains a new note (with the complementary probability the agent does not enter an institution). Finally, consider \( v^{n+1}_0 (m) \). At the end of a period, an agent without a note expects to enter an institution with probability \( \kappa \), in which case she obtains no notes or goods (with the complementary probability the agent does not enter an institution).

Proposition 2 establishes conditions under which the candidate steady state exists.

**Proposition 2** Let \( \beta \geq \max \left\{ \frac{1 - c}{\kappa u + mc}, 1 - m \right\} \) and

\[
nm \left( 1 - m \right) (u - c) \geq a + \left( 1 - \beta \right) m \left( (1 - m) u + mc \right).
\]
Then, there exists a steady state equilibrium in which the measure of agents with notes equals \( m \) and agents and managers behave consistently with \( P = (1, 1, 0, 1, 0) \).

We have already discussed how the visibility of banks prevents them from overissuing notes. Proposition 2 demonstrates that as long as agents are sufficiently patient and in each period the frequency of meetings in which exchange occurs (as measured by \( nm (1 - m) \)) is sufficiently large, notes circulate. In particular, inspection of the conditions in the proposition reveals a critical property of the bank-based equilibrium. Unlike in the case of a fair-based equilibrium, such an equilibrium exists regardless of the relative size of institutions, as captured by \( \kappa \), as long as \( \kappa \) is positive and \( \beta \) is sufficiently high. The economic intuition is simple. The value of notes has two components: their circulation value (as a medium of exchange) 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 redeem notes with a low probability), this can always be offset by a higher circulation value, which in turn can be achieved by a sufficiently large amount of notes in circulation. Put differently, even when institutions become relatively sparse and, hence, less valuable for agents in redeeming notes, they can preserve an essential role by costlessly issuing notes that serve as a medium of exchange (i.e., injecting liquidity).

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 gift-exchange allows trades to occur in all private meetings while note-exchange does not so because it requires that in a meeting one agent has a note and her mate has zero notes.

### 3.4 From Fairs to Banks

In this section, we demonstrate that the economy can endogenously transit from a state in which institutions act as fairs - and trade is sustained by a norm of gift-exchange - to a state in which institutions act as banks - and trade is sustained by the circulation of notes issued by managers. The force that drives this transition is the expansion of trade. The rationale for the transition is contained in Propositions 1 and 2: the former proves that gift-exchange cannot be sustained when the size of institutions relative to the population involved in trade is small, whereas the latter proves that a note-exchange equilibrium exists regardless of the relative size of institutions. We preserve the environment adopted so far but we allow for changes in the extent of trade by assuming that, besides the original unit continuum of agents, the economy also comprises a large number of “newcomers”. At the beginning of every period, a measure \( \Omega \) of newcomers is born. Each newcomer is productive with probability \( \pi \approx 1 \), and her type is her private information. At the beginning of period \( t \), institutions host a measure \( \omega_t \in [0, \Omega] \) of newly born newcomers as well as a measure 1 of agents or newcomers who have already engaged in exchange, randomly chosen from their respective populations. We let \( \omega_1 = 0 \), that is in the first period all agents but no newly born newcomers enter
institutions. To ease exposition, henceforth we call “agents” newcomers who have already engage in exchange.

As a preliminary point, observe that newcomers never exchange with each other because there are no institutions sustaining their trade. Moreover, in the absence of notes, if agents were unable to distinguish between productive and unproductive newcomers, agents would not be willing to engage in private meetings with newcomers. In fact, in such meetings all unproductive newcomers would announce “consume” because they are unable to produce. This implies that a productive newcomer would have the incentive to misrepresent her type and also announce “consume”. By doing so, she would increase her current payoff and preserve the same future payoff because her deviation would have zero measure and her identity would not be publicly observable. However, our economy allows to overcome this problem. In fact, inside institutions transactions are publicly observable and, by producing for the manager of an institution, a productive newcomer reveals her type, credibly separating herself from unproductive ones.\(^5\) This makes clear the crucial role that the visibility of institutions plays in integrating new traders into the exchange process.

The above reasoning implies that the measure of agents at the beginning of period \(t\) equals

\[ \eta_t = \eta_{t-1} + \pi \omega_{t-1}, \]

that is the measure of agents in the previous period plus the newly born productive newcomers who have exchanged inside institutions at the beginning of period \(t\) (all the unproductive newcomers and the previously born productive newcomers who did not enter the exchange process will stay in autarky). The measure \(\eta_t\) of agents weakly increases over time so that the probability \(\frac{1}{\eta_t}\) that a randomly chosen agent enters an institution weakly decreases over time.

We now demonstrate that this in turn eventually leads to the collapse of fair-based exchange and the transition to bank-based exchange. Here we provide an informal presentation of the result while the formal proof is relegated to the Appendix.

In Section 3.2, we proved that gift-exchange can be sustained on the equilibrium path if and only if (for all \(i \in \{0, \ldots, n - 1\}\))

\[ \frac{\beta}{1-\beta} (v_f - v_d) \geq \frac{\beta}{1-\beta} (1-\kappa)^{2^{n-i}} (v_{2^{n-i}} - v_d) + \frac{1}{2} (n-i) (u+c) + \frac{\beta \kappa}{2} (3u-c). \] \hspace{1cm} (10)

Remember that in (10) \(v_d\) is the expected payoff at the beginning of private meetings right after a deviation is observed inside institutions. As explained in Section 3.2, after such a deviation is observed, an agent stays in autarky in the current and in all future periods (which gives a payoff of \(a\)) and continues to exchange inside institutions in all future periods (which gives an expected payoff of \(\beta \frac{3}{2} \kappa (u-c)\)). As a result, \(v_d \equiv a + \beta \frac{3}{2} \kappa (u-c)\). In this section, we propose a strategy such that, if an agent has observed a deviation inside an institution, she stops cooperating, stays in autarky in the current period and, in all future periods, she only produces in exchange for notes. This implies that \(v_d\) is replaced by \(v(m) \equiv (1-\beta) a + \beta w(m)\), where \(w(m)\) is the expected payoff in a steady

\(^5\)We are implicitly assuming that, once a productive newcomer has revealed her type, she can always be recognized as such in her private meetings. We can think that the manager assigns tokens to productive newcomers.
state where notes circulate and the measure of agents with notes is $m$. A second modification of (10) stems from the fact that the expected payoff from gift-exchange changes over time because now the probability $\frac{1}{\eta_t}$ that an agent enters an institution changes. Therefore, we need to replace $v_f$ with $v_{f,t+1}$, where

$$v_{f,t+1} = \frac{1}{2} n (u-c) + \frac{3}{2} (u-c) \sum_{s=0}^{\infty} (1-\beta) \frac{1}{\eta_{t+1+s}}.$$  \hspace{1cm} (11)

All in all, taking these two changes into account, gift-exchange can be sustained on the equilibrium path in period $t$ if and only if (for all $i \in \{0,\ldots,n-1\}$)

$$\frac{\beta}{1-\beta} [v_{f,t+1} - v(m)] \geq \frac{\beta}{1-\beta} \left( 1 - \frac{1}{\eta_{t+1}} \right)^{2^{n-i}} [v_{2^{n-i},t+1} - v(m)] + B_i(\beta,\eta_{t+1}),$$  \hspace{1cm} (12)

where $B_i(\beta,\eta) \equiv \frac{1}{2} (n-i) (u+c) + \frac{\beta}{2\eta} (3u-c)$. In condition (12), $v_{2^{n-i},t+1}$ is the expected payoff at the beginning of private meetings in period $t+1$ when no deviation was observed inside institutions and the number of agents who are not cooperating equals $2^{n-i}$. The computation of $v_{2^{n-i},t+1}$ is similar 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 at the beginning of their first private meeting, her expected flow payoff from following the norm is $(1-\beta)nu$. At the beginning of the following period, there

is a probability $\left( 1 - \frac{1}{\eta_{t+2}} \right)^{2^{n-i}}$ that a deviation will not be observed inside institutions, in which case the expected continuation payoff is $v_{2^{n-i},t+2}$. 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 $(1-\beta) \frac{\eta_{t+3}u}{1-(1-\frac{1}{\eta_{t+2}})2^{n-i}} u + v_m$. Thus, we obtain

$$v_{2^{n-i},t+1} - v(m) = (1-\beta) \sum_{s=0}^{\infty} \beta^s \prod_{j=2}^{s+1} \left( 1 - \frac{1}{\eta_{t+j}} \right)^{2^{n-i}} \left[ nu - v(m) + \beta \frac{1}{\eta_{t+s+2}} u \right].$$  \hspace{1cm} (13)

Using (13), condition (12) can be rewritten as (for all $i \in \{0,\ldots,n-1\}$)

$$\frac{\beta}{1-\beta} [v_{f,t+1} - v(m)] \geq \sum_{s=0}^{\infty} \beta^{s+1} \prod_{j=1}^{s+1} \left( 1 - \frac{1}{\eta_{t+j}} \right)^{2^{n-i}} \left[ nu - v(m) + \beta \frac{1}{\eta_{t+s+2}} u \right] + B_i(\beta,\eta_{t+1}).$$  \hspace{1cm} (14)

Let us inspect (14) to grasp the intuition behind the transition from fair-based exchange to bank-based exchange. Under gift-exchange, production and consumption occur in all meetings, while under note-exchange production in a meeting requires that one agent has a note while her mate has zero notes. Hence, it must be the case that $v_{f,t+1} > v(m)$. This implies that, if $\{\eta_t\}_{t=1}^{\infty}$ is bounded, there exists $\beta(n,\{\eta_t\}_{t=1}^{\infty})$ large enough that (14) is satisfied. Intuitively, if agents meet relatively often inside institutions, as long as they are sufficiently patient, gift-exchange can be sustained on

\[\text{We define } \prod_{j=2}^{\infty} \left( 1 - \frac{1}{\eta_{t+j}} \right)^{2^{n-i}} \equiv 1.\]
the equilibrium path. In contrast, condition (14) implies that if \( \{\eta_t\}_{t=1}^{\infty} \) is unbounded there will exist a period \( t(n, \beta, \{\eta_t\}_{t=1}^{\infty}) \) at and after which it is violated. We can see this by letting \( \eta_t \to \infty \), which implies \( \frac{1}{\eta_t} \to 0 \) in (14). Intuitively, at any level of patience, if \( \frac{1}{\eta_t} \) converges to zero, there exists a period when an agent has the incentive to unilaterally deviate from gift-exchange because she knows that anyway the contagion process initiated by her deviation will reach an institution - and, hence, backfire on her - far in the future. Hence, the expansion of exchange eventually leads to the collapse of fair-based exchange and the transition to bank-based exchange.

A formal proof that a transition from fair-based exchange to bank-based exchange indeed occurs as an equilibrium phenomenon is not trivial. The reason is that, if agents take into account that gift-exchange will be abandoned in period \( t(n, \beta, \{\eta_t\}_{t=1}^{\infty}) \), they will have the incentive to stop cooperating in period \( t(n, \beta, \{\eta_t\}_{t=1}^{\infty}) - 1 \). In turn, if agents stop cooperating in period \( t(n, \beta, \{\eta_t\}_{t=1}^{\infty}) - 1 \), no agent will have the incentive to cooperate in period \( t(n, \beta, \{\eta_t\}_{t=1}^{\infty}) - 2 \). Proceeding in this fashion, gift-exchange would collapse in period 1, when the relative size of fairs is large and a small number of agents engage in exchange. We avoid this unraveling of cooperation by considering a strategy profile that punishes deviations inside institutions that occur before \( t(n, \beta, \{\eta_t\}_{t=1}^{\infty}) \) with a reversion to autarky, while it punishes deviations inside institutions that occur after \( t(n, \beta, \{\eta_t\}_{t=1}^{\infty}) \) with a reversion to note-exchange. Since note-exchange provides a higher payoff than autarky, agents have the incentive to cooperate in private meetings up to period \( t(n, \beta, \{\eta_t\}_{t=1}^{\infty}) \) despite the fact that the economy will transit to note-exchange.

Proposition 3 summarizes our result. We say that bank-notes are “essential” if they achieve desirable allocations that could not be achieved otherwise.

**Proposition 3** There exists an equilibrium such that, up to some finite period \( t(n, \beta, \{\eta_t\}_{t=1}^{\infty}) \) fair-based exchange is sustained and bank-notes are not essential while after that bank-notes become essential. From period \( t(n, \beta, \{\eta_t\}_{t=1}^{\infty}) \) onwards, the economy is in a steady-state equilibrium in which agents and managers behave according to the vector \( P = (1, 1, 0, 1, 0) \) described in Proposition. In addition, in this steady-state, newly born newcomers always produce for managers and receive a note with probability \( m \). This equilibrium Pareto dominates any other equilibrium.

An interesting side result of our analysis is that fairs generate the conditions for their own demise. It is because fairs are visible that they allow to reveal which new exchangers will comply with the norm of gift-exchange, allowing to integrate them into the exchange process. Yet, it is exactly this expansion of exchange that eventually undermines the role of fairs and requires the transition to banks. Of course, the reader could think of alternative reasons for which exchange expands, such as an exogenous process of population growth. Here, we have preferred focusing on an endogenous process of exchange expansion driven by the visibility of institutions.
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 Renaissance Europe. It is worth stressing up-front that our aim is not to obtain hard evidence but simply to verify whether the historical facts are broadly consistent with our hypothesis. We first evaluate whether our key assumption - the nature of fairs and banks as visible institutions - is grounded in the account of historians. Next, we discuss the more traditional view of fairs as “primitive courts”. Finally, we turn to historical evidence on the transition from fair-based trade to bank-based trade. We evaluate whether banks and the associated credit instruments indeed emerged as the evolution of fairs and trade instruments and whether this evolution stemmed from the progressive inability of fairs to sustain the expansion of trade.

There was public information on transactions in Medieval and Renaissance fairs and banks.

There is historical evidence that a large amount of public information and records was available about transactions occurring in Medieval fairs. In the Cambridge Economic History of Europe, Verlinden (1965, p. 128) reports that “It is therefore certain that there existed at the Champagne fairs a real records department” and “a meagre fragment has recently been found: a leaf from a register of the Troyes summer fair of 1296. Drawn up by an Italian notary, it contains fifteen deeds [...] These deeds are concerned with exchange transactions [...] Seven of the fifteen are letters by which a carrier undertakes to convey to their destination within a prescribed time consignments of merchandise”. Moreover, Verlinden (1965, p. 137) reports that 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.” Similarly, there is also historical evidence on the large amount of public information that was available about transactions occurring in banks in Medieval and Renaissance Europe. In particular, the books of banks were publicly available for inspections and, hence, transactions could easily be detected. De Roover (1948, p. 265) 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.” De Roover (1948, p. 265) also argues that though it is “open to doubt whether in Bruges the journals enjoyed recognition as public records [...] whether or not such was the case makes no practical difference. In case of need the money changer could always be called upon to testify [...] with regard to any transaction which he had [...] entered into his books.” He also adds (p. 266) that “the public character of bank records made such a formality [to require a voucher for each deposit] superfluous.”

Fairs could not sustain an efficient system of 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 an aspect of the role of fairs. In fact, the ability to enforce promises in distant trades was very limited in Medieval Europe: “Because of the slowness of communication the control of agents in distant places remained one of the knotty problems of mercantile capitalism until the end of the eighteenth century. 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).” The experience of the Venetian merchant Guglielmo Querini is illuminating. “This merchant spent twelve years in futile and obstinate attempts to collect outstanding claims in Flanders, in England and in the Levant. In only one instance did he succeed, but it was near home and not in distant lands (De Roover, 1965, p.88).” The problem of securing satisfactory representation [...] is also exemplified by the career of another Venetian merchant, Andrea Barbarigo, who had his share of troubles with unreliable correspondents, too. De Roover (1965, p.89) summarizes these problems as follows: “Medieval business letters give the impression that principals were often disappointed because their agents sold their consignments for less than they had expected to get or paid too high a price for local commodities.” Even when fairs carried out some form of legal enforcement, this practice took time to spread and was often summary. According to Verlinden (1965, p. 131) “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.” Moreover, Verlinden (1965, pp. 136-37) argues that “in Flanders [the exercise of justice and the procedure in their courts] was usually summary. Deeds of agreement often contained a waiver of legal process likely to slow up the course of the proceedings.”

In Medieval Europe, as trade expanded, fairs evolved towards financial institutions.

The progressive evolution from fairs to banks is well documented by historians. Verlinden (1965, pp. 132-33) writes that “From the middle of the thirteenth century onwards, money-changing begins to take precedence over trade. By 1262 the letters of the Tolomei of Siena provide evidence of this transformation [...] The Tolomei in France obtained at the fairs the necessary funds for their purchases of cloth by means of advances, loans, exchanges made by their Italian head office. For purchases at a fair the agent went north at the time of the preceding fair and the Tolomei funds he employed were considered as an advance, his account being debited with the interest on this sum from one fair to another. The Del Bene of Florence, or even earlier, the Sienese in 1294, operated a similar system [...]. In any case, from this time 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. Currency and bill quotations at the fairs, which quickly became known outside through the fair couriers, became a factor in international speculation. In fact the financial system of the Lions fairs in the fifteenth century, the Besancon fairs in the sixteenth and those of Piacenza up to the
seventeenth, was already heralded in Champagne by the end of the thirteenth century [...]” And also (Verlinden, 1965, p. 133): “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 historical evidence also suggests that the evolution of fairs was triggered by the expansion of trade (which was promoted by fairs themselves). This is especially evident if one looks at the type of instruments that were used 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, as the ius mercatorum gradually recognized the validity of informal instruments (De Roover, 1965, p. 69).”

5 Conclusion

In this paper, we have put forward a theory for the origin of banks based on the role of information in trade. We have studied a decentralized economy where information on transactions is scarce. We have shown that in such an economy trade can initially be sustained by institutions where transactions are observed and recorded publicly. We have interpreted these institutions as fairs. We have then demonstrated that the expansion of trade allowed by fairs can progressively undermine their ability to disseminate information about non-cooperative behaviors. When trade becomes intense, institutions can continue to sustain it only if they start issuing notes and acting as banks. The sustainability of bank-based trade hinges on the same feature of institutions that supported fair-based exchange, that is their visibility, but is not undermined by the expansion of trade. We have argued that the model can help rationalize the decline of fairs and the emergence of banks in Medieval and Renaissance Europe.

References


6 Appendix

Proof of Proposition 1: First, we need to show that, for any given \( \kappa \), there exists a unique \( \beta(\kappa) \) such that (for \( i \in \{0, ..., n - 1\} \))

\[
\frac{\beta \left[ \frac{1}{2} n (u - c) - a \right]}{1 - \beta} \geq (nu + \beta \kappa c - a) \sum_{s=0}^{\infty} \beta^{s+1} (1 - \kappa) \frac{2^{n-i}(2^n+2^n-2)}{2^{n-1}} + \frac{1}{2} (n - i) (u + c) + \beta \kappa \quad (15)
\]

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 (15) converges to zero when \( \beta \) converges to zero and converges to infinity when \( \beta \) converges to one. In turn, the right-hand side (RHS) of (15) converges to \( \frac{1}{2} (n - i) (u + c) \) when \( \beta \) converges to zero and converges to

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

when \( \beta \) converges to one. This implies that there exists \( \beta_i(\kappa) \) such that (15) holds with the equality sign. Moreover, the first derivative of the RHS of (15) with respect to \( \beta \) is

\[
\partial \text{RHS} \quad \frac{\partial}{\partial \beta} = \left\{ \begin{array}{l} \kappa c + (nu - a) \sum_{s=0}^{\infty} (s+1) \beta^s (1 - \kappa) \frac{2^{n-i}(2^n+2^n-2)}{2^{n-1}} + \\ \kappa c \sum_{s=0}^{\infty} (s+2) \beta^{s+1} (1 - \kappa) \frac{2^{n-i}(2^n+2^n-2)}{2^{n-1}} \end{array} \right\} > 0, \quad (17)
\]

and the second derivative of the RHS of (15) with respect to \( \beta \) is

\[
\partial \left( \frac{\partial \text{RHS}}{\partial \beta} \right) = \left\{ \begin{array}{l} (nu - a) \sum_{s=0}^{\infty} s (s+1) \beta^{s-1} (1 - \kappa) \frac{2^{n-i}(2^n+2^n-2)}{2^{n-1}} + \\ \kappa c \sum_{s=0}^{\infty} (s+1) (s+2) \beta^s (1 - \kappa) \frac{2^{n-i}(2^n+2^n-2)}{2^{n-1}} \end{array} \right\} > 0. \quad (18)
\]

This implies that the RHS of (15) is strictly convex in \( \beta \). Since the LHS of (15) is also strictly convex in \( \beta \), there exists a unique \( \beta_i(\kappa) \) that satisfies (15) with the equality sign. Finally, let

\[
\beta(\kappa) = \max_{i \in \{0, ..., n - 1\}} \beta_i(\kappa).
\]

We now prove that (15) 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 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 \( \frac{3}{2} \kappa (u - c) + a \). Therefore, the agent will follow the norm if and only if

\[
\frac{1}{2} n (u - c) \geq a + (1 - \beta) \frac{1}{2} (u + c).
\]

(19)

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

\[
\frac{1}{2} n (u - c) \geq a + (1 - \beta) c.
\]

(20)
Note that
\[
\frac{1}{2}n(u - c) \geq a + \frac{1}{\beta} \left[ (1 - \beta) (nu + \beta kc - a) \sum_{s=0}^{\infty} \beta^{s+1} (1 - \kappa) \frac{2^{s-1} (2^{s+2} - 2)}{2^{s-1}} \right],
\]
which corresponds to (15), satisfies (19) and (20).

We now prove ii and iii (a). If she has observed a deviation in a private meeting but no deviation inside institutions (a) an agent participates in exchange but she does not cooperate in private meetings, (b) an agent deviates in her first meeting with a manager. Moreover, if she has observed deviations inside institutions once, an agent stays in autarky but she continues to cooperate inside institutions. The strongest incentive to slow down the spread of non-cooperative behavior occurs when (1) the agent engages in a meeting inside 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 rule, she deviates and her expected payoff is \( (1 - \beta) \left[ u + \frac{1}{2} (u - c) + a \right] + \beta \left[ \frac{1}{2} \kappa (u - c) + a \right]. \) 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, in which case her expected payoff is \( (1 - \beta) \frac{3}{2} (u - c) + \beta \frac{3}{2} \kappa (u - c) + a \). There is a complementary probability that such a deviation will not occur and her expected payoff is \( (1 - \beta) \frac{3}{2} (u - c) + v_2 \). Comparing her expected payoffs, the agent will follow the norm and deviate if and only if
\[
(1 - \beta) c \geq (1 - \kappa) (v_2 - v_d),
\]
which can be rewritten as
\[
\sum_{s=0}^{\infty} \beta^s (1 - \kappa) \frac{2 (2^{s-1} + 2^{s-1})}{2^{s-1}} \leq \frac{c}{nu + \beta kc - a}.
\]
Fix \( \beta \in (0, 1) \). The left-hand side (LHS) of (23) converges to \( \frac{c}{nu-a} \) when \( \kappa \) converges to zero and converges to zero when \( \kappa \) converges to one. Moreover, the LHS of (23) is strictly decreasing in \( \kappa \). In turn, the right-hand side (RHS) of (23) is strictly decreasing in \( \kappa \), it converges to \( \frac{c}{nu-a} \) when \( \kappa \) converges to zero, and it is equal to \( \frac{c}{nu+\beta kc-a} \) when \( \kappa \) is equal to one. Because \( \frac{c}{nu-a} < 1 \leq \frac{1}{\beta} \), there exists a unique \( \kappa^*(\beta) \) such that (23) holds for all \( \kappa \geq \kappa^*(\beta) \). Note that a sufficient condition for (23) to hold, which occurs when \( \beta \) converges to 1, is
\[
\sum_{s=0}^{\infty} (1 - \kappa) \frac{2 (2^{s-1} + 2^{s-1})}{2^{s-1}} \leq \frac{c}{nu + \beta kc - a}.
\]
Finally, we also need to prove iii (b) if she has observed deviations inside institutions more than once, an agent stays in autarky and she does not cooperate with defectors inside institutions. First, assume that a deviation inside institutions has already occurred in period \( t \), and consider the incentives to deviate inside an institution in some period \( t' > t \). If all agents follow the norm, in period \( t' \) there is no cooperation outside institutions but there is cooperation inside them. If an agent cooperates with a manager, she obtains an expected payoff \( (1 - \beta) \frac{3}{2} (u - c) + \beta \frac{3}{2} \kappa (u - c) + a \). If she deviates,
she obtains \((1 - \beta) u + \beta \kappa (u - c) + a\) because she will be excluded from all future production in her meetings with other agents inside institutions. The agent will not deviate if and only if

\[
\beta \geq \frac{3c - u}{3c - u + \kappa (u - c)}.
\]

(25)

An agent may also have an incentive to deviate in a meeting with another agent. If she does not deviate, she obtains \((1 - \beta) \frac{1}{2} (u - c) + \beta \frac{2}{3} \kappa (u - c) + a\), while a deviation implies an expected payoff of \((1 - \beta) u + \beta \kappa (u - c) + a\). The agent will not deviate if and only if

\[
\beta \geq \frac{u + c}{u + c + \kappa (u - c)}.
\]

(26)

Clearly, (26) implies (25). Hence, for any given \(\kappa\), there exists a unique \(\beta' (\kappa)\) such that (26) holds if and only if \(\beta \geq \beta' (\kappa)\). Henceforth, define \(\beta^* (\kappa) = \max \{ \beta (\kappa), \beta' (\kappa) \}\). Summarizing, for any given \(\beta\) and \(\kappa\), there exists \(\beta^* (\kappa)\) and \(\kappa^* (\beta)\) such that the norm of gift exchange is an equilibrium if and only if \(\beta \geq \beta^* (\kappa)\) and \(\kappa \geq \kappa^* (\beta)\). In particular, since a sufficient condition for (23) to hold occurs when \(\beta\) converges to 1, the norm of gift exchange is an equilibrium for all \(\beta \geq \beta^* (\kappa)\) and \(\kappa \geq \lim_{\beta \to 1} \kappa^* (\beta)\).

**Proof of Proposition 2:** We prove that there exists a steady state equilibrium with the following features: (i) agents always produce in exchange for notes \((\lambda = 1)\), (ii) notes are always redeemed \((\gamma = 1)\), (iii) agents without notes do not receive any good from a manager \((\eta = 0)\), (iv) managers offer new notes in exchange for a good to all agents who have just redeemed an old note, and only to them \((\alpha_1 = 1\) and \(\alpha_0 = 0)\), (v) agents always participate in private exchange. Moreover, if any manager deviates from \(P\) no agent will produce for this manager in any future meeting and no agent will produce in exchange for notes issued by this manager in any future meeting. Before proving our claim, it is useful to display again the value functions implied by the candidate steady-state \(P = (1, 1, 0, 1, 0)\) (for \(i \in \{0, \ldots, n\}\), where \(i = 0\) corresponds to the meeting inside an institution)

\[
v^i_1 (m) = me_1^{i+1} (m) + (1 - m) [u + v^{i+1}_1 (m)],
\]

(27)

\[
v^0_0 (m) = m [-c + v^{i+1}_1 (m)] + (1 - m) v^{i+1}_0 (m),
\]

(28)

\[
v^{n+1}_0 (m) = \beta [\kappa v^0_0 (m) + (1 - \kappa) v^1_0 (m)],
\]

(29)

and

\[
v^{n+1}_1 (m) = \beta \{ \kappa [u - c + v^0_0 (m)] + (1 - \kappa) v^1_0 (m) \}.
\]

(30)

First, an agent is willing to produce in exchange for a note as long as \(-c + v^1_1 (m) \geq v^0_0 (m)\), for all \(i \in \{0, \ldots, n\}\), and \(-c + v^{n+1}_1 (m) \geq v^{n+1}_0 (m)\). The first inequality always holds, while the second holds as long as

\[
\beta \geq \frac{c}{c + (1 - m) (u - c) + \kappa (u - c)}.
\]

(31)

We also need to check under what conditions an agent prefers participating in exchange than staying in autarky. Clearly, \(-c + v^1_1 (m) \geq v^0_0 (m)\) implies that it is always better to participate in exchange
with a note. Hence, we only need to check the incentives of an agent without a note. From (27) and (28), the expected payoff of participating in exchange without a note is

\[ v_0^1(m) = v_0^n(m) + (n - 1) m (1 - m) (u - c), \]

where

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

An agent without a note prefers participating in exchange than staying in autarky if and only if

\[ v_0^1(m) \geq a + v_0^{n+1}(m), \]

that is

\[ -mc + (n - 1 + \beta) m (1 - m) (u - c) + \beta m [\kappa (u - c) + c] \geq a. \]

It remains to be shown that each manager is always willing to behave according to \( P = (1, 1, 0, 1, 0) \). The flow payoff of a manager who follows this strategy equals the measure \( \kappa m \) of goods that she stores each period. Her best possible deviation is to offer notes in exchange for goods to all agents in her institution, in which case her flow payoff is \( (1 - \beta) \kappa \) and her continuation payoff is zero because no agent will ever produce for this manager again. Comparing payoffs, the manager will not deviate if and only if

\[ \beta > 1 - m. \]

Clearly, no agent has an incentive to produce for the notes of this manager because these notes are no longer accepted as a medium of exchange. Summarizing, there exists a steady state equilibrium in which the measure of agents with notes equals \( m \) and agents and managers behave consistently with \( P = (1, 1, 0, 1, 0) \) if and only if

\[ \beta \geq \max \left\{ \frac{c}{c + (1 - m)(u - c) + \kappa (u - c)}, 1 - m \right\} \]

and

\[ -mc + (n - 1 + \beta) m (1 - m) (u - c) + \beta m [\kappa (u - c) + c] \geq a. \]

We now turn to demonstrate that there exists a transition from an initial state in which no agent holds a note to a steady state where a measure \( m \) of agents hold a note. The argument runs as follows. Assume that, up to period \( t - 1 \), all agents who meet a manager without a note receive a note. In turn, all agents who meet a manager without a note at the beginning of period \( t \) receive a note with probability \( q \). Assume also that, up to period \( t \), there is no redemption of notes and agents do not need to produce in exchange for a note. This implies that, at the beginning of period \( t \), the measure of agents with notes is given by

\[ m_t = \kappa + (1 - \kappa) \kappa + ... + (1 - \kappa)^{t-1} q \kappa. \]

We can rewrite this expression as

\[ m_t = 1 - (1 - \kappa)^{t-1} (1 - q \kappa). \]
For all \( q \in [0, 1] \), \( \frac{\partial m}{\partial t} > 0 \) 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

\[
m = 1 - (1 - \kappa)^{t_m - 1} (1 - q_m \kappa). \tag{40}
\]

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

\[
m_t = \begin{cases} 
1 - (1 - \kappa)^t, & \text{if } t < t_m \\
1 - (1 - \kappa)^{t_m - 1} (1 - q_m \kappa) = m, & \text{if } t = t_m \\
m, & \text{if } t > t_m
\end{cases} \tag{41}
\]

This path converges to \( m \) in a finite number of periods \( t_m \). As stated above, along the transition path, agents do not produce for managers and notes are unredeemable. This guarantees that, even if an agent is not planning to stay in autarky, she will accept a note because there is no cost in doing so and she may use the note in the future. In turn, managers are visible and can be convinced to issue notes without receiving goods in exchange as long as they expect some future benefit (punishment) associated with following (not following) this course of action. Now, consider the economy in period \( t_m + 1 \), when the measure of agents with a note equals \( m \). At \( t_m + 1 \), even though there are no goods in storage, managers can redeem “old” notes, offer no good in exchange for notes, and ask agents to produce in case they want to receive a new note. Again, because meetings between agents and managers are visible, an agent with an old note can be convinced to give it back to the manager as long as she is properly punished (rewarded) in case she keeps (does not keep) her old note. This implies that, starting in period \( t_m + 2 \), notes can always be redeemed in exchange for goods. From this point onwards, the economy is in a steady-state where, as long as (31) and (34) hold, notes are valued as a medium of exchange and the amount of notes in circulation equals \( m \). Moreover, managers and agents obtain a positive expected payoff, which provides the necessary compensation for their behavior along the transition path. Precisely, if either a manager or an agent deviates from the proposed behavior along the transition path, all agents can coordinate on never producing inside institutions and always staying in autarky. Since both managers and agents expect to obtain positive payoffs in the future, they will not deviate.

Proof of Proposition 3: We first describe the strategies of agents, newcomers and managers. Henceforth, we say that an agent cooperates if she announces “produce” and that a manager cooperates if she uniformly distributes her stored goods to all agents inside her institution.

Agents: The history of an agent includes the actions in all her past private meetings, and the actions of agents, newcomers, and managers in all past meetings inside institutions. An agent cooperates at the beginning of period 1.

Consider the history of cooperation in private meetings and meetings inside institutions that occur up to period \( t^* \). After this history, (a) if the agent is making decisions before period \( t^* \), she cooperates in all meetings and participates in private meetings; (b) if she is making decisions during period \( t^* \), she cooperates inside institutions and stays in autarky (if she is in a private meeting, she deviates); (c) if she is making decisions after period \( t^* \), she does not cooperate but she produces in exchange for notes if and only if the
economy is in a steady-state where notes are valued and there is a fraction \( m \) of agents with notes. If the economy is not in this steady-state, she deviates in all meetings and stays in autarky (if she is in a private meeting, she deviates).

Consider histories of cooperation inside institutions and deviations in private meetings that occur up to period \( t^* \). After these histories, if the agent is making decisions before period \( t^* \), she cooperates inside institutions, participates in private meetings and deviates in these meetings. If the agent is making decisions during period \( t^* \), she cooperates with the manager, she deviates in her meeting with another agent inside the institution, and stays in autarky (if she is in a private meeting, she deviates). If the agent is making decisions after period \( t^* \), she behaves in exactly the same way as after the cooperative history.

Finally, consider histories of deviations inside institutions that occur up to period \( t^* \). In this case, if a deviation happened before the meeting between agents inside institutions in period \( t^* \), an agent deviates in all meetings and stays in autarky (if she is in a private meeting, she deviates). If, instead, deviations occurred for the first time in the meeting between agents inside institutions in period \( t^* \), (i) each agent responsible for the deviation does not cooperate with the manager, cooperates in her meeting with another agent, and stays in autarky (if she is in a private meeting, she deviates), (ii) each agent not responsible for this deviation cooperates inside institutions and stays in autarky (if she is in a private meeting, she deviates). If any agent deviates from either (i) or (ii), she deviates in all meetings and stays in autarky (if she is in a private meeting, she deviates).

**Newcomers**: We assume that, upon entering an institution, newcomers have the same history as managers, that is, the past history of actions inside institutions. Conditional on her history, productive newcomers behave in the same way as agents. In turn, newcomers who did not participate inside institutions stay in autarky in all periods.

**Managers**: In all periods \( t \leq t^* \), a manager cooperates. In all periods \( t > t^* \), if the economy is in a steady-state in which notes circulate as a medium of exchange, and there is a fraction \( m \) of agents with notes, managers do not cooperate but they (a) redeem old notes, (b) issue new notes in exchange for goods to all agents who have just redeemed an old note, and (c) issue notes in exchange for goods to a random fraction \( m \) of newcomers. If the economy is not in this steady state, managers cooperate.\(^7\)

We first establish the conditions under which there are no deviations from the equilibrium path. In a period \( t > t^* \), the economy is in a steady-state in which notes circulate as a medium of exchange and the fraction of agents with notes equals \( m \). This scenario is essentially the same as that described in Section 3.3, the only difference being that the probability that an agent enters an institution evolves according to \( \frac{1}{\eta_t} \). As proved in Section 3.3., as long as \( -mc + (n - 1)m(1 - m)(u - c) \geq a \) and \( \beta \geq \beta_{agent}(m) \equiv \frac{c}{c + (1 - m)(u - c)} \), an agent has an incentive to produce in exchange for a note, regardless of the value of \( \frac{1}{\eta_t} \). It remains to be shown that each manager is always willing to follow her proposed strategy. The flow payoff of a manager who follows her strategy equals the measure \( m \) of goods that she stores each period. Her best possible deviation is to offer notes in exchange for goods that she stores each period. Her best possible deviation is to offer notes in exchange for goods that she stores each period.

\(^7\)We also assume that, at the beginning of period 1, managers randomly distribute notes to a fraction \( m \) of agents. Moreover, from period 1 to period \( t^* \), managers distribute notes to a fraction \( m \) of all newcomers that are able to produce. This ensures that, when notes start to circulate as a medium of exchange, there is a fraction \( m \) of agents holding notes. Finally, from period 1 to period \( t^* - 1 \), managers distribute tokens to all newcomers that are able to produce. This ensures that, in private meetings, productive newcomers can be recognized as such.
goods to all agents in her institution, in which case her flow payoff is 1 − β and her continuation payoff is 0 because agents and newcomers never cooperate again. We obtain that a manager will not deviate if and only if \( \beta \geq \beta_{\text{manager}}(m) \equiv 1 - m \).

Consider now the economy in period \( t^* \). Clearly, because all agents choose to stay in autarky during this period, it is optimal to stay in autarky. Inside institutions, if an agent wants to deviate from the equilibrium path, it is best to deviate against another agent (as opposed to deviate against a manager) because it provides the same flow payoff and a higher continuation payoff. Moreover, because it is always better to enter note-exchange with a note, an agent without a note has the highest incentive to deviate in period \( t^* \). This reasoning implies that a sufficient condition under which no deviation occurs inside institutions in period \( t^* \) is

\[
(1 - \beta) \left\{ \frac{1}{2} (u - c) + a \right\} + \beta \nu_1^1(m) \geq (1 - \beta) u + \beta \left[ u + \frac{1}{2} (u - c) \right] + a,
\]

where \((1 - \beta) \nu_1^1(m) = -1 - \beta mc + (n - 1 + \beta) m (1 - m) (u - c)\) is a lower bound on the agent’s expected payoff of entering note-exchange without a note at the beginning of period \( t^* + 1 \). After some manipulation, we can rewrite (42) as

\[
\frac{\beta}{1 - \beta} \left[ -(1 - \beta) mc + (n - 1 + \beta) m (1 - m) (u - c) - \left( a + \frac{3}{2} u - \frac{1}{2} c \right) \right] \geq \frac{1}{2} (u + c).
\]

Clearly, as long as \( n > \frac{a + \frac{3}{2} u - \frac{1}{2} c + m[(1 - m) u + mc]}{m[1 - m](u - c)} \), there exists \( \beta(n) \) such that (43) holds for all \( \beta \geq \beta(n) \). In turn, because a manager faces the same problem as in period \( t > t^* \), she does not deviate if and only \( \beta \geq \beta_{\text{manager}}(m) \).

We also need to establish the conditions under which there are no deviations from the equilibrium path in a period \( t < t^* \). Consider an agent in the \( i^{th} \) private meeting in period \( t \), where \( i \in \{1, ..., n\} \). The highest possible payoff that this agent can obtain after a deviation occurs when her deviation does not reach an institution up to the meetings between agents in period \( t^* \). In this case, the number of agents that are reached by this deviation (including the agent herself) is \( 2^{(t^* - t)n + 1 - i} \). Taking this particular realization into account, we obtain that a sufficient condition ensuring cooperation along
the equilibrium path is
\[
(1 - \beta) (n - i + 1) \frac{1}{2} (u - c) + \beta \left(1 - \beta^{t^* - t - 1}\right) \frac{1}{2} (u - c) + (1 - \beta) \beta^{t^* - t} \frac{1}{\eta_t} \frac{3}{2} (u - c) + \beta^{t^* - t + 1} (1 - \beta) \frac{1}{\eta_t} (m) \geq (1 - \beta) (n - i + 1) u + \beta \left(1 - \beta^{t^* - t - 1}\right) nu + (1 - \beta) \beta^{t^* - t} \frac{1}{\eta_t} (2u - c) + \beta^{t^* - t + 1} \left(1 - \frac{1}{\eta_t}\right) 2^{(t^* - t) + 1 - i} (1 - \beta) \frac{1}{\eta_t} (m) + \beta^{t^* - t + 1} \left[1 - \left(1 - \frac{1}{\eta_t}\right) 2^{(t^* - t) + 1 - i}\right] \left[u + \frac{1}{2} (u - c) + a\right].
\]

A sufficient condition for (44) is
\[
\frac{\beta^{t^* - t + 1}}{(1 - \beta^{t^* - t})} \frac{1}{\eta_t} (n) + (1 - \beta) \beta^{t^* - t} \frac{1}{\eta_t} \left[1 - \left(1 - \frac{1}{\eta_t}\right) 2^{(t^* - t) + 1 - i}\right] \text{LHS} \ (43) \geq \frac{1}{2} (u + c) \quad (45)
\]
Clearly, as long as \( n > \frac{a + (u - c) + \frac{1}{2} (u + c) + m (1 - m) u + mc}{\eta_t} \), there exists \( \beta \ (n, \eta_t, t^*) \) such that (45) holds for all \( \beta \geq \beta \ (n, \eta_t, t^*) \). \(^9\) Note that whenever (45) holds, (43) also holds. Now consider the behavior of an agent inside an institution triggers an immediate collapse of cooperation and a transition to autarky, if the agent does not deviate in private meetings, she does not deviate inside institutions. Finally, \( \beta > 1 - m \) ensures that managers do not deviate in any period \( t < t^* \). In sum, as long as
\[
n > \frac{a + (u - c) + \frac{1}{2} (u + c) + m [(1 - m) u + mc]}{\eta_t} (1 - m) (u - c), \quad (46)
\]
and
\[
\beta \geq \max \left\{ \beta \_\text{agent} \ (m), \beta \_\text{manager} \ (m), \beta \ (n, \eta_t, t^*) \right\}, \quad (47)
\]
there will be no deviations along the equilibrium path.

We now establish the conditions under which there are no deviations from the proposed strategies after any out-of-equilibrium-path history. We first look at histories of cooperation inside institutions and deviations in private meetings that occur up to period \( t^* \). An upper bound on the expected payoff that an agent can obtain after a deviation inside an institution in a period \( t < t^* \) is \((1 - \beta) 2u + \beta (1 - \beta) u + \alpha\), because no agent ever cooperates again and all agents stay in autarky. \(^{10}\)

\(^9\) Clearly, \( \beta \ (n, \eta_t, t^*) \) is increasing in \( t^* \) reflecting the fact that agents need to be more patient if the transition to note-exchange only happens far in the future. In particular, if the transition never occurs so that \( t^* \) converges to infinity, condition (47) cannot be satisfied.

\(^{10}\) \((1 - \beta) 2u\) is an upper bound on the flow payoff inside an institution after the deviation, and \( \beta (1 - \beta) u \) corresponds to the possibility that the agent still receives one unit of good from a manager in period \( t + 1 \).
In turn, a lower bound on the expected payoff that the agent can obtain by following the proposed strategy is \((1 - \beta^{t^* - 1}) nu + \beta^{t^* - 1} a\). Intuitively, the agent deviates in all private meetings up to period \(t^* - 1\) and receives at least the autarky payoff from period \(t^*\) on. A sufficient condition ensuring that the agent follows the proposed strategy is

\[
(1 - \beta^{t^* - 1}) nu + \beta^{t^* - 1} a \geq (1 - \beta) (2u + \beta) u + a.
\] (48)

A sufficient condition for (48) is \(n \geq \frac{a + 3u}{2}\). Now consider the agent’s decision inside an institution in period \(t^*\) after histories of cooperation inside institutions and deviations in private meetings that occur up to period \(t^*\). Clearly, the agent has no incentive to deviate in her meeting with a manager because she can obtain the same flow payoff and a higher continuation payoff by cooperating with the manager and (as proposed by the strategy) deviating with another agent inside the institution. Hence, we only need to consider a deviation from the proposed strategy in her meeting with another agent. If the agent follows the proposed strategy she deviates in her meeting with another agent and stays in autarky. Her expected payoff in this case is

\[
(1 - \beta) u + \beta \left( \frac{3}{2} u - \frac{1}{2} c \right) + a.
\] (49)

An agent may be tempted to deviate from the proposed strategy and cooperate in her meeting with another agent inside the institution. By behaving this way, there is a positive probability that no other agent will deviate inside the institution in period \(t^*\), in which case the economy will transit to note exchange and will not revert to autarky. The agent’s temptation to cooperate depends on her belief about the number of agents who are going to deviate in meetings with other agents inside the institution. In turn, under our proposed strategy, this number equals the number of agents who also observed a deviation in some private meeting up to period \(t^*\). For a given strategy profile, beliefs are trivially computed by using Bayes rule along the path of play induced by the strategy profile. However, Bayes rule cannot be applied to out of the equilibrium path-histories. In order to deal with this issue, we resort to the trembling-hand theory. Moreover, because trembles by a finite number of agents are infinitely more likely than trembles by a continuum of agents, we postulate that, upon observing a deviation in a private meeting, an agent must infer that the deviation was caused by and is restricted to a finite number of agents.\(^{11}\) Consistent with this interpretation, we assume that, after observing a deviation in a private meeting for the first time, the agent infers that this deviation occurred in the first private meeting in period 1. As a result, she believes that the number of agents (excluding herself) reached by a deviation at the beginning of period \(t^*\) is at least \(2^{(t^* - 1)n - 1}\). Conditional on this belief, an upper bound on the expected payoff of cooperating in the

\(^{11}\)A similar reasoning is put forth by Takahashi (2007). Takahashi investigates whether a community with a continuum of agents supports cooperation in the repeated prisoner’s dilemma when agents are anonymously and randomly matched in pairs.
meeting with another agent inside an institution in period $t^*$ is

$$(1 - \beta) \frac{1}{2} (u - c + a) + \beta \left( 1 - \frac{1}{\eta_{t^*}} \right)^{2^{(t^* - 1)n - 1}} (1 - \beta) \pi_1^t (m) + \beta \left[ 1 - \left( 1 - \frac{1}{\eta_{t^*}} \right)^{2^{(t^* - 1)n - 1}} \right] \left[ \frac{3}{2} (u - c) + a \right],$$

where $(1 - \beta) \pi_1^t (m) = u - mc + (n + 1) m (1 - m) (u - c)$ is an upper bound on the agent’s expected payoff of entering note exchange with a note at the beginning of period $t^* + 1$. An agent follows the proposed strategy as long as

$$\beta \left( 1 - \frac{1}{\eta_{t^*}} \right)^{2^{(t^* - 1)n - 1}} \left[ c + (1 - \beta) \pi_1^t (m) - \left( \frac{3}{2} u - \frac{1}{2} c + a \right) \right] \leq (1 - \beta) \frac{1}{2} (u + c) + \beta c. \quad (50)$$

We can rewrite this expression as

$$\beta \left( 1 - \frac{1}{\eta_{t^*}} \right)^{2^{(t^* - 1)n - 1}} \left\{ (1 - m) c + (n + 1) m (1 - m) (u - c) \right\} \leq (1 - \beta) \frac{1}{2} (u + c) + \beta c. \quad (51)$$

It is sufficient to evaluate (52) at $\beta = 1$. We obtain

$$\left( 1 - \frac{1}{\eta_{t^*}} \right)^{2^{(t^* - 1)n - 1}} \left\{ (1 - m) c + (n + 1) m (1 - m) (u - c) - \left[ a + \frac{1}{2} (u - c) \right] \right\} \leq c. \quad (53)$$

For all pair $(\eta_{t^*}, t^*)$, there exists $n (\eta_{t^*}, t^*)$ sufficiently large such that $n \geq n (\eta_{t^*}, t^*)$ satisfies (53).

So far, we have been considering only the agent’s decision inside institutions in a period $t \leq t^*$, after observing a deviation in a private meeting but observing no deviations inside institutions up to period $t^*$. Clearly, if the agent has no incentive to deviate from the proposed strategy by cooperating in her meeting with another agent inside the institution in period $t^*$, she has no incentive to slow down the contagion process in private meetings up to period $t^*$. Moreover, since all agents stay in autarky in period $t^*$, the agent also stays in autarky and does not participate in private exchange in this period. Finally, we need to consider the agent’s decision in a period $t > t^*$, after histories of cooperation inside institutions and deviations in private meetings that occur up to period $t^*$. Since the agent’s behavior in this case is identical to the behavior after the cooperative history, we have already proved above that the agent has no incentive to deviate.

Consider now histories of deviations inside institutions that occur up to period $t^*$. It may be the case that the deviation inside an institution occurred for the first time in the meeting between agents in period $t^*$. After this history the proposed strategy postulates that (i) each agent responsible for the deviation does not cooperate with the manager, cooperates in her meeting with another agent, and stays in autarky (if she is in a private meeting, she deviates), (ii) each agent not responsible for this deviation cooperates inside institutions and stays in autarky. This behavior induces a positive expected payoff to all agents that is higher than the autarky payoff. As a result, and because any
deviation from the proposed behavior would lead to no cooperation in all meetings, no agent has an incentive to deviate. Finally, it may also be the case that a deviation happened before the meeting between agents inside institutions in period $t^*$. In this case, the proposed strategy asks for a deviation by all agents in all future meetings and for agents to stay in autarky. Clearly, any deviation from this behavior implies a reduction in the flow payoff and no change in the continuation payoff.\footnote{Note that the same reasoning ensures that no deviation from the proposed strategy will occur after a history that contains a deviation from the note-exchange equilibrium where the fraction of agents with notes is equal to $m$.}

There is no need to check the incentives of managers in the absence of a note-exchange equilibrium. The reason is that, after any such histories, a deviation by a manager from the proposed strategy does not affect her flow payoff and it may decrease her continuation payoff. Moreover, there is no need to check the incentives of a newcomer who participates inside institutions because she has the same strategy as an agent. Finally, because no agent will cooperate with a newcomer who does not enter an institution, such newcomers always stay in autarky.

In summary, we have shown that, given a pair $(\eta_{t^*}, t^*)$, as long as

$$n \geq \max \left\{ \frac{n(\eta_{t^*}, t^*)}{\frac{a + (u - c) + \frac{1}{m} (u + c) + m [(1 - m) u + mc]}{m (1 - m) (u - c) \frac{1}{u}} + \frac{a + 3 u}{u}} \right\},$$

and $\beta \geq \max \{\beta_{agent}(m), \beta_{manager}(m), \beta(n, \eta_{t^*}, t^*)\}$, there is an equilibrium under which, in period $t^*$, the economy transits from gift-exchange to note-exchange.

As a last step, we need to relate the period $t^*$ in which the economy transits from gift-exchange to note-exchange to the period in which agents have an incentive to deviate from gift-exchange, regardless of the behavior of other agents. This ensures that the transition to note-exchange is not driven by a coordination motive, where agents abandon one scheme of exchange and move to another scheme simply because all other agents are doing the same. Precisely, we need to make sure that the entry dynamics of newly born newcomers $\{\omega_t\}_{t=1}^{\infty}$ and thus the induced dynamics of the measure of agents $\{\eta_t\}_{t=1}^{\infty}$ is such that period $t^*$ is the first period when

$$\frac{\beta}{1 - \beta} [v_{f, t^* + 1} - v(m)] < \sum_{s=0}^{\infty} \beta^{s+1} \prod_{j=1}^{s+1} \left( 1 - \frac{1}{\eta_{t^* + j}} \right) \left( nu - v(m) + \beta \frac{1}{\eta_{t^* + s + 2}} \right) + B_i(\beta, \eta_{t^* + 1}).$$

Given $(\eta_{t^*}, t^*)$, $\beta$ and $n$, we can always find a decreasing sequence $\{\eta_t\}_{t=1}^{\infty}$ that satisfies this requirement.