# Dynamic Formation of Credit Networks
# CS 261: Final Report

Ana Carolina Queiroz, Sahasrajit Sarmasarkar, Hoden Abdi-Yusuf

June 2022

**Abstract**

Current efforts to understand how credit networks are strategically formed analyze one-shot games, in which trust is assigned once. We propose a dynamic framework with Bayesian agents in which trust evolves based on prior beliefs of the overall trustworthiness of the network, and information acquire through interactions between two agents. We propose two models. In a model in which agents have private beliefs, agents who have lower credit-issuing limits overestimate the trustworthiness of the network, whereas agents with larger limits have beliefs that converge faster, and to a value closer to the real network trustworthiness. In a model in which agents have shared beliefs, after some nodes initially default, agents underestimate the trustworthiness of the network for several rounds.

# 1 Introduction

Credit networks are a useful abstraction to model trust in a network of individuals [1, 2]. Transactions between two individuals $u$ and $v$ who don't trust each other directly can occur as long as there is a directed path of users who trust each other, connecting $u$ and $v$ [2].

Our project focuses on the strategic formation of credit networks. In their seminal work on the topic, Dandekar et al. [2] study a one-shot game in which self-interested agents belonging to an exogenous network $H$ decide how much trust they will assign to other agents in the resulting network $G$. We build on their work to study network formation with repeated games, in which agents have Bayesian beliefs about the *trustworthiness* of other agents in the network.

We adopt an externalist position of trust: the person who assigns trust to someone else has rationally estimated this amount, and updates this belief as new information becomes available [3]. From this perspective, it becomes important to investigate how trust networks arise in a dynamic setting, in which beliefs about the trustworthiness of other agents can evolve.

The rest of the paper is organized as follows: In Section 2, we develop our main model of dynamic formation of credit networks and document the notation. Section 3 presents our methods and an alternative specification in which agents have shared beliefs about each other's trustworthiness. Section 4 contains our main results and a brief discussion. Section 5 concludes and outlines potential directions for future work.

# 2 A Model of Dynamic Formation

We set up the dynamic model of strategic formation of credit networks as follows: We have an initial network $H = (V, E)$, with agents partitioned into two types, trustworthy ($T$) and untrustworthy ($U$). There are $\mu \cdot |V|$ trustworthy agents and $(1 - \mu) \cdot |V|$ untrustworthy agents. The true $\mu$ is not observed by any of the agents, but it is common knowledge that $\mu$ is drawn from a uniform distribution $\mu \sim Unif[0, 1]$. Each agent also knows their true type.

The following game is played repeatedly. In the first step of the $i^{th}$ game, as in [2] each node $u$ decides how much credit (trust) to assign to each of the other nodes, and is also bounded by some credit limit $B_u$ it can extend to others. This process creates a graph $G^{(i)}$. We then allow for a series of transactions to occur in $G^{(i)}$, in which IOUs are exchanged between neighbors. In the second step, each node $u$ "reclaims" for each neighbor $v$ the IOUs issued by $v$ that are in $u$'s possession at the end of the transaction stage. At this point, each node $u$ decides whether to default and not honor the IOUs it issued for $v$ (i.e. $d^{(i)}_{(u,v)} = 1$) or to honor the IOUs (i.e. $d^{(i)}_{(u,v)} = 0$).

Each step below further specifies the game played in each iteration $i$:

**Step 0**

We assume that agents are Bayesian. The beliefs of agent $u$ towards the trustworthiness of an agent $v$ in round $i$, $\tilde{\mu}^{(i)}_{(u,v)}$, is a function of $u$'s belief about the overall trustworthiness of all agents in the network, $\hat{\mu}^{(i)}_{(u,V)}$, and the past behaviors of $v$ that have been observed by $u$, $d_{(v,u)}$ (i.e. whether a node has defaulted or not). After one round of transactions in a network $G^{(i)}$, $u$ updates their belief about $v$ based on their new belief of the overall trustworthiness of the network and on the newly-observed behavior of $v$. Our goal is to incorporate positive and negative reinforcement [4] on the beliefs agent $u$ has about the other agents: A positive interaction with $v$ (i.e. $v$ honor IOUs issued to $u$) marginally increases $\hat{\mu}^{(i+1)}_{(u,V)}$ (belief of overall trustworthiness in network for the net round) and $\tilde{\mu}^{(i+1)}_{(u,v)}$ (belief of trustworthiness of node $v$ for the next round). A negative interaction decreases both values, and we assume that once a node $v$ defaults on $u$, $u$ never extends credit to $v$ again. Even if one node $z$ did not default on a transaction with $u$, the belief of $u$ towards $z$'s trustworthiness might decrease if $u$ has other negative interactions in previous rounds, since $\tilde{\mu}^{(i)}_{(u,z)}$ is also a function of the overall belief $\hat{\mu}^{(i)}_{(u,V)}$.

**Updates on overall belief:** At the beginning of each iteration $i$, the updated overall belief of $u$ about the trustworthiness of the entire network, $\hat{\mu}^{(i)}_{(u,V)}$, is the average of the belief of trustworthiness $u$ has of each other individual in the network in the previous iteration $i-1$ (the average of all $\tilde{\mu}^{(i-1)}_{(u,v)}$):

$$\hat{\mu}^{(i)}_{(u,V)} = \frac{\sum\limits_{v:(u,v)\in E} \tilde{\mu}^{(i-1)}_{(u,v)}}{|V|-1}$$

**Updates on beliefs towards individuals:** At the end of each iteration $i$, the updated belief that $u$ has on the trustworthiness node $v$, $\tilde{\mu}_{(u,v)}$, is the weighted average of the overall belief and the past interactions between $u$ and $v$ (the term is always 1 if $v$ does not default on $u$ up to iteration $i$). :

$$\tilde{\mu}^{(i)}_{(u,v)} = \begin{cases} \alpha_i \cdot \hat{\mu}^{(i)}_{(u,V)} + (1-\alpha_i)\cdot 1 \text{ if } v \text{ did not default on } u \text{ up to iteration } i \\ 0 \text{ otherwise} \end{cases}$$

Note that $\alpha_i$ is a term that controls how fast individual interactions with $v$ become more important to $u$ than the overall beliefs $u$ holds about the trustworthiness of the entire network. For our analysis, we set $\alpha_i$ to be $\alpha_i = \left(\frac{1}{i+1}\right)^{0.2}$. Note that the larger the exponent (which we set to 0.2), the faster the beliefs will assign greater weight to individual interactions in a round and lower weight to overall beliefs.

**Step 1**

In the first step of each round, agent $u$ decide whether to form a tie with a given $v$, and if so, how much credit to extend to $v$, $c^{(i)}_{(u,v)}$, such that it maximizes its expected gains from trade $g_u(.)$ minus the losses $l(.)$ from potential defaults of neighbors. Note that we define $g_u(.)$ (gain of node $u$ from trading) as a function of the credit extended by various nodes. The losses are a function of the belief $u$ has of each $v$ defaulting, given the amount of credit $u$ extends to $v$ and the (estimated) trustworthiness of $v$. We make the simplifying assumption that trade can only happen between neighbors. The variables in bold are vectors with the corresponding values for each $v$, for a given $u$. The utility for $u$ function given a network $G^{(i)}$, $\gamma_u^{(i)}(G^{(i)})$, is given by:

$$\gamma_u^{(i)}(G^{(i)}) = \mathbb{E}\left[g_u(\{\boldsymbol{c^{(i)}_{(u,v)}}\}_{v:(u,v)\in E})\right] - \sum_{v:(u,v)\in E} \mathbb{E}\left[l(\boldsymbol{c^{(i)}_{(u,v)}}, \boldsymbol{\tilde{\mu}^{(i)}_{(u,v)}})\right]$$

We set the gains-of-trade function to $g_u(.) = \sum_{v:(u,v)\in E}(2 \cdot \sqrt{c^{(i)}_{(u,v)} + 1} - 2)$ and the expected loss function to be $l(.) = (1 - \mu^{(i)}_{(u,v)})c^{(i)}_{(u,v)}$. The gains-of-trade function captures the idea of diminishing returns for each extra unit of credit (and trade) extended to a certain node. The loss function ensures that when a node has defaulted and revealed to $u$ it is untrustworthy, the ideal credit $u$ amount will issue is going to be zero. We further assume that all agents in iteration $i$ try to maximize their utility, given the credit issuing limit $B_u$ that each node $u$ has. This boils down the the convex maximization program:

$$\max \sum_{v:(u,v)\in E} (2 \cdot \sqrt{c^{(i)}_{(u,v)} + 1} - 2) - \sum_{v:(u,v)\in E} (1 - \mu^{(i)}_{(u,v)}) \cdot c^{(i)}_{(u,v)}$$

Subject to $u$'s own credit-issuing constraint, $B_u$:

$$\sum_{v:(u,v)\in E} c^{(i)}_{(u,v)} \leq B_u$$

**Steps 2 and 3**

In the second step, given the credit network, the neighbors transact with each other, up to the credit amount that has been extended to them. In the third step, each $u$ decides whether to default or honor a transaction with $v$. Note that a node $u$ can only observe whether $v$ has defaulted or honored a transaction if they are neighbors and have transacted. Recall we have two types of agents: trustworthy ($T$) and untrustworthy ($U$). If $u$ is of type $T$, $u$ will always honor a transaction.

As a simplifying assumption, for $u$ of type $U$, defaulting is contingent on whether the amount of credit $u$ issued to $v$ ($c^{(i)}_{(u,v)}$) is greater than the amount of credit $v$ issued to $u$ ($c^{(i)}_{(v,u)}$). If this statement is true ($c^{(i)}_{(u,v)} \geq c^{(i)}_{(v,u)}$), then we assume $u$ does not have an incentive to default on $v$, since $v$ owes more units to $u$ than $u$ owes units to $v$. Note that we assume each node transacts up to the amount of credit that is issued to them. Otherwise, $u$ defaults, with some probability that is a product of $\left(\frac{c^{(i)}_{(v,u)} - c^{(i)}_{(u,v)}}{\sum_{v\in V\{u\}} \max\{c^{(i)}_{(v,u)} - c^{(i)}_{(u,v)}, 0\}}\right)$ and a fixed default probability $\delta$. Intuitively, we would like the probability of defaulting to be higher if the gap between how much $u$ owes to $v$ and how much $v$ owes to $u$ is the greatest, as a share of all the profitable potential defaults $u$ could make in round $i$. Therefore, the default indicator function is:

$$d_{u,v}^{(i)} = \begin{cases} 0 & \text{if } u \in T \text{ or } c_{(u,v)}^{(i)} \geq c_{(v,u)}^{(i)} \\ 1 & \text{with probability } p = \delta \cdot \left( \frac{c_{(v,u)}^{(i)} - c_{(u,v)}^{(i)}}{\sum_{v \in V} \max\{c_{(v,u)}^{(i)} - c_{(u,v)}^{(i)}, 0\}} \right) \end{cases}$$

# 3 Methods

Our main goal is to determine if (1) the beliefs of agents in the network converge and (2) if the network becomes stable, that is, if there is some some $c$, such that for all iterations $t > c$ the set of nodes with which a node $u$ transacts remains the same. We obtain preliminary results using simulations.

For the simulations, we randomly generate $|V| = 40$ agents with credit issue constraint $B_u \sim Unif[1, 100]$ once. We set the initial $\hat{\mu}_{(u,V)}^{(0)} = 0.5$ for all $u \in V$, which is the expected value of a random variable $X \sim Unif[0, 1]$. We set the default probability $\delta = 0.3$. We also randomly assign each node to the trustworthy or untrustworthy set with probability $\mu \in \{0.25, 0.5, 0.75\}$, simulating each state of $\mu$ separately. In each iteration, each of the nodes perform the convex optimization from step 1. For each specification of $\mu$, we run 400 rounds of the game.

We also perform simulations for an alternative model, in which beliefs about the trustworthiness of a node are shared across all other nodes. In this model, the beliefs all nodes have over the trustworthiness of a node $v$ is:

$$\tilde{\mu}_v^{(i)} = \begin{cases} \alpha_i \cdot (\hat{\mu}^{(i)}) + (1 - \alpha_i) \cdot 1 \text{ if node } v \text{ has not defaulted on any node up to round } i \\ 0 \text{ otherwise} \end{cases}$$

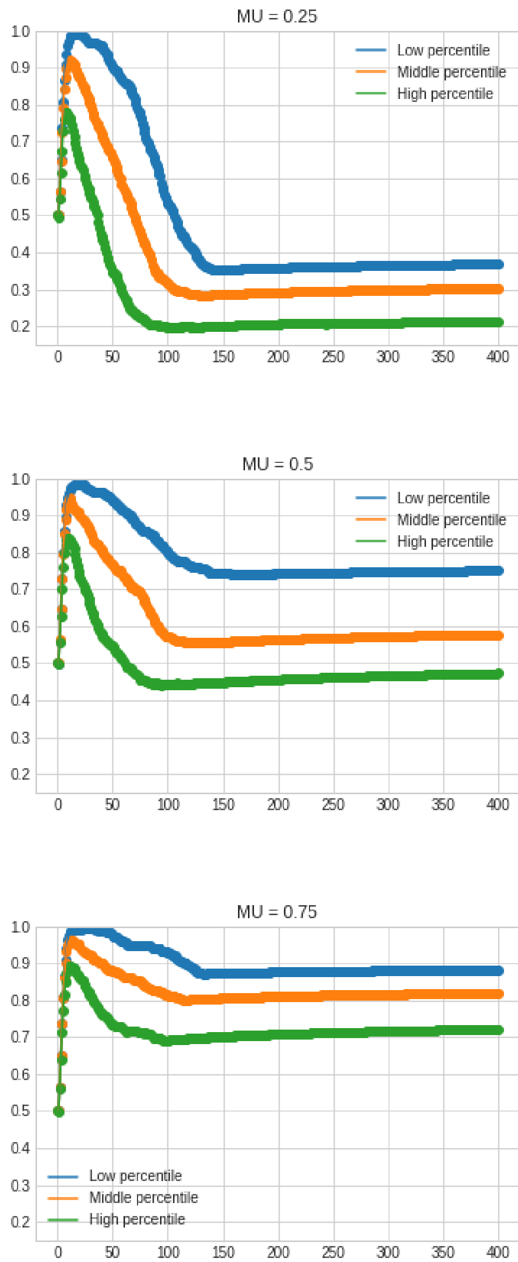
Such that $\hat{\mu}^{(i)}$ is the overall beliefs that all nodes have of the trustworthiness of the overall network. This belief is updated at the beginning of each $i$, in the same way we update overall beliefs $\hat{\mu}_{(u,V)}^{(i)}$ on the private-information case. All other steps remain the same in this alternative model.

# 4 Results and discussion

In Figure 1, we plot the evolution of the average of the overall beliefs $\mu_{(u,v)}^{(i)}$ of network trustworthiness for the agents on the different terciles of credit-issuing values $B_u$ (i.e. high, middle, and low percentiles). Each plot represents the evolution under a different true value of network trustworthiness, $\mu = 0.25, \mu = 0.50, \mu = 0.75$. Recall that higher values of $\mu$ imply that more agents will never default.

For all three $\mu$ specifications, the terciles behaved similarly in relation to each other. The average network trustworthiness beliefs of agents on the lower tercile (i.e. those who could extend the least credit to others

Figure 1: Estimated Network Trustworthiness, by Credit-Issuing Limit Terciles for different True Network Trustworthiness $\mu$

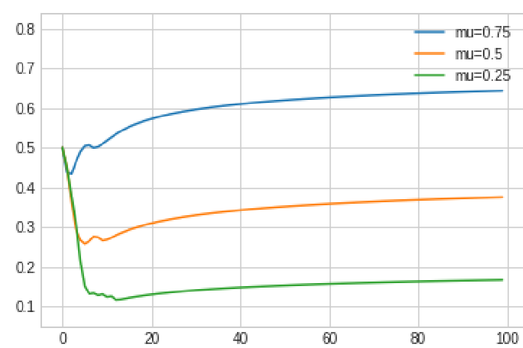



Note: The $y$ axis represents the average network trustworthiness beliefs of agents on the high, middle, and low terciles of the credit-issuing limit distribution on a given round $i$, $\mu^{(i)}_{(u,V)}$. The $x$ axis represents the round in which the updated belief is generated. Each panel uses the same set of agents, with different underlying network trustworthiness $\mu = 0.25, 0.50, 0.75$.

in the network) overestimated the true $\mu$ in all cases. The middle and high tercile agents had beliefs that better approximated the true $\mu$, with agents on the middle tercile slightly overestimating the trustworthiness of the network in all cases. The beliefs seem to converge on all cases, with faster convergence for the agents on the high tercile. This is likely due to the larger amount of credit extended to each edge it assigns trust to: The larger the amount of credit an untrustworthy agent can get, the larger is the likelihood that they will default and reveal their true type at any given round. For agents who cannot extend large amount of credit, there is less incentive for untrustworthy agents to default, so the defaults are more spread across time, and the overall beliefs take longer to update, eventually converging in some overestimate of the trustworthiness of the network. We also note that across the different $\mu$ values, all groups converge to some belief, so, under our model, we cannot claim that more or less trustworthiness in the network determines whether the beliefs will converge.

The convergent beliefs on all cases are a good indicator that the underlying networks becomes more stable over time: The set of agents with which one is willing to transact with narrows down, and their beliefs about these individual agents becomes more stable, leading to a more stable overall belief of the network trustworthiness.

Figure 2: Estimated Network Trustworthiness (Shared Belief)
by different True Network Trustworthiness $\mu$




Note: The $y$ axis represents the shared network trustworthiness beliefs on a given round $i$, $\hat{\mu}^{(i)}$. The $x$ axis represents the round in which the updated belief is generated. Each panel uses the same set of agents, with different underlying network trustworthiness $\mu = 0.25, 0.50, 0.75$.

In Figure 2, we consider the same exercise, but with shared beliefs as described in Section 3. In this case, all agents have the same overall belief of network trustworthiness in each round $i$. With our specification, the beliefs start to converge before the $20^{th}$ round. It is interesting to observe that for the case in which half of the network is trustworthy, there is a drop on overall beliefs before this value increases. Within the first 100 iterations, it is still below the true $\mu$. The same pattern happens for a lower and higher value of $\mu$ ($\{0.25, 0.75\}$). This indicates that when a node defaults in a network with shared beliefs under our specifications, it takes many iterations for the agents to recover the trust on the network.

## 5 Conclusion and Next Steps

We have created a dynamic model in which agents strategically form credit networks, updating their beliefs about the trustworthiness of individual nodes and the entire network in each round. We find that the amount of credit extended by a node $u$ is important to determine whether other nodes will default on $u$, which shapes how fast $u$'s beliefs about trustworthiness will converge, and whether they will converge to the correct trustworthiness parameter. Nodes that can extend more credit are more susceptible to defaults, and learn more quickly. We also find that, under our specification, overall belief seem to converge for high, medium, and low levels of network trustworthiness. In a shared-beliefs model, we also observe that defaults deteriorate the overall belief of network trustworthiness, which is underestimated for several rounds.

Some future directions for work could be exploring random fluctuations in the initial credit assigned, as well as modelling how nodes could strategize forming edges in cases in which transactions in paths (and not just between two neighbors) is allowed.

## References


[1] Pranav Dandekar et al. "Liquidity in credit networks: a little trust goes a long way". In: *Proceedings of the 12th ACM conference on Electronic commerce*. EC '11. New York, NY, USA: Association for Computing Machinery, June 2011, pp. 147–156. ISBN: 978-1-4503-0261-6. DOI: 10.1145/1993574.1993597. URL: https://doi.org/10.1145/1993574.1993597 (visited on 05/13/2022).

[2] Pranav Dandekar et al. "Strategic Formation of Credit Networks". In: *ACM Transactions on Internet Technology* 15.1 (Mar. 2015), 3:1–3:41. ISSN: 1533-5399. DOI: 10.1145/2700058. URL: https://doi.org/10.1145/2700058 (visited on 05/11/2022).

[3] Carolyn McLeod. "Trust". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2021. Metaphysics Research Lab, Stanford University, 2021. URL: https://plato.stanford.edu/archives/fall2021/entriesrust/ (visited on 05/11/2022).

[4] Brian Skyrms and Robin Pemantle. "A Dynamic Model of Social Network Formation". en. In: *Adaptive Networks: Theory, Models and Applications*. Ed. by Thilo Gross and Hiroki Sayama. Understanding Complex Systems. Berlin, Heidelberg: Springer, 2009, pp. 231–251. ISBN: 978-3-642-01284-6. DOI: 10.1007/978-3-642-01284-6_11. URL: https://doi.org/10.1007/978-3-642-01284-6_11 (visited on 05/13/2022).

# A Appendix: Table of Notation

| Term | Definition |
|---|---|
| $d^{(i)}_{(u,v)}$ | Node $u$ honours IOU to node $v$ at round $i$ |
| $\tilde{\mu}^{(i)}_{(u,v)}$ | Belief of trustworthiness of node $u$ about node $v$ at round $i$ |
| $\hat{\mu}^{(i)}_{(u,V)}$ | Belief of trustworthiness of node $u$ about all nodes $V$ in graph $G^{(i)}$ at round $i$ |
| $c^{(i)}_{(u,v)}$ | Credit extended to node $v$ by node $u$ at round $i$ |
| $B_u$ | Credit issuing limit of node $u$ |
| $x^{(i)}_{(u,v)}$ | IOU's that node $u$ owes node $v$ at round $i$ |
| $\alpha_i$ | weight of overall beliefs *vs.* individual interactions to set individual beliefs in round $i$ |
| $\delta$ | Overall default probability of an un-trustworthy agent. |

Table 1: Summary of Notation

# B Appendix: Permanent links to replication code

Dynamic Formation Simulations (Individual Beliefs)

Dynamic Formation Simulations (Shared Beliefs)