Substitutability in Favor Exchange

Oğuzhan Çelebi*

June 3, 2024

Abstract

I study a favor exchange model in which players enforce cooperation bilaterally and can rely on multiple partners for favors, that is, relationships are potentially substitutable. With substitutability, the frequency players interact and the value of their relationships are determined by the network, and the equilibrium exhibits empirically observed intermediate cooperation. With heterogeneous players, substitutability causes homophily and exacerbates inequality. Community enforcement prevents bilateral ties and cannot be combined with bilateral enforcement. By considering substitutability, my model can explain the stratification of social networks in post-Soviet states and the absence of bilateral relationships for medieval traders who practiced community enforcement.

*Department of Economics, Stanford University, 579 Jane Stanford Way, Stanford, CA 94305. Email: ocelebi@stanford.edu.

I am grateful to Daron Acemoglu, Yunus Aybas, Daniel Barron, Jonathan Bendor, Roberto Corrao, Glenn Ellison, Joel Flynn, Drew Fudenberg, Robert Gibbons, Ali Kakhbod, Giacomo Lanzani, Juan Ortner, Parag Pathak, Heikki Rantakari, Karthik Sastry, Alp Simsek, Birger Wernerfelt, Alex Wolitzky and participants in the MIT Theory Lunch, MIT Sloan Organizational Economics Lunch and Stanford Political Economy Theory Lunch for helpful comments and suggestions. First Draft: November 2021.

1. Introduction

In many settings, ranging from villages in rural India to trade agreements in medieval Europe, individuals and firms depend on informal mechanisms to sustain cooperation. In these settings, writing down and enforcing explicit contracts is hard due to high costs or the lack of infrastructure, and cooperative behavior enforced by informal mechanisms is important. The enforcement of cooperation has been studied extensively, usually focusing on how coalitions of individuals can punish those who behave in uncooperative ways and the role that the network of relationships plays in enforcing cooperation.[1] This paper focuses on another role of the network; the role it plays in determining how and when players interact with each other, which in turn determines the values they derive from their relationships.

I study a model in which players linked on a network exchange costly favors with their partners over time, which extends the earlier models in multiple dimensions. First, instead of assuming any two linked players play with an exogeneously given probability, I decompose the process that determines when players play to two: the need for a favor and the ability to provide a favor. Second, I allow for different types of favors, where each player may have a different probability of being able to provide each favor type. Two important special cases of this model are *monopolistic favors*, where each player has a unique favor type such that only they can provide, and *substitutable favors* where players are homogeneous in terms of their favor provision abilities. Substitutable favors are natural in many settings; examples include lending small amounts of money, providing childcare to a neighbor, or covering a shift for a coworker, all of which can possibly be provided by many individuals in their respective settings. Monopolistic favors correspond to more specific cases where individuals have clear comparative advantage, such as lending a valuable equipment that no one else owns or sharing a niche expertise.

I show that assuming any two linked players play with an exogeneously given probability is equivalent to monopolistic favors in my setting. This is because, when favors are monopolistic, two players play when one requires the favor type the other player can provide, and the other player is able to provide a favor in that period. As both of these probabilities do not depend on the rest of the network, the probability that they interact and the benefit they receive from their relationships do not depend on the rest of their network. I show that these features cause monopolistic models to make unreasonable predictions of no cooperation or universal cooperation when relationships are enforced bilaterally, explaining the earlier focus on community (multilateral) enforcement.

[1] See Wolitzky (2021) for a recent survey of this literature, as well as other important early contributions from other disciplines such as Granovetter (1985) and Coleman (1994).

In contrast, when favors are substitutable, the network structure plays an important role in determining the frequency of interactions and the values players derive from their relationships, exhibiting several more realistic and empirically relevant features. First, the equilibrium cooperation in the society is intermediate and varies continuously with the primitives, instead of jumping from no cooperation to universal cooperation. Second, when players are heterogeneous, equilibrium networks exhibit stratification, where high type players (*e.g.*, players who can provide favors more easily) have larger networks that consists of other high type players. This prediction is aligned with empirical observations and explains how social networks play an important role in exacerbating inequality. Third, community enforcement prevents establishment of bilateral ties and the cannot be combined to with bilateral enforcement to enhance cooperation. This result explains why certain groups that practiced community enforcement have not established bilateral relationships with outsiders.

Substitutability of Favors and Value of Relationships. Substitutability of favors is not only more realistic in many settings, but also leads to two mechanisms that cause the network structure to determine the value each player derives from their relationships and are central to the results of this paper. First, when favors are substitutable, the probability that any two players interact in any given period, and thus the benefit they receive from their relationships, depend critically on the network structure. In particular, an additional relationship affects the provision of a given favor when it is pivotal, that is, that player is the only player who can provide that favor in that period. The probability that any given partner is pivotal, and hence the marginal value of a relationship decreases in the number of relationships. Second, players prefer to have relationships with those who have a high number of relationships, as they require favors less frequently, and players with a low number of relationships are undesirable partners.

Bilateral Enforcement. Most of the research on social cooperation focus on community enforcement, where the network structure plays an important role in enforcing cooperation. However, in many settings, the relationships between players are enforced bilaterally, without any sanctions or punishment from other players.[2] Moreover, studying bilateral enforcement

[2]Under bilateral enforcement, the action of player i when playing with player j can only depend on the previous interactions between i and j, but not on the previous interactions with others. Bilateral mechanisms play an important role in enforcing cooperation in many settings. For example, due to the lack of markets, favor exchange (*blat*) relationships were crucial for obtaining necessary goods and services in the Soviet Union and were mostly enforced through bilateral reciprocity (see Section 4.2). In the medieval era, traders from different cultural backgrounds formed different enforcement mechanisms. Mahgribi traders relied on a multilateral enforcement mechanism, while Genoese traders relied on a legal system and a bilateral reputation mechanism (see Section 5.1). Moreover, it is reasonable to think that in modern societies where social networks are used to obtain professional favors, the relationships are enforced bilaterally due to the individualistic culture and lack of transparency and monitoring compared to village societies where

allows us to abstract away from the role network plays in enforcing cooperation and focus on its role in determining the value created by the relationships. I first study the benchmark case of monopolistic favors and show that under bilateral enforcement, there is a sharp and unrealistic discontinuity in the predicted levels of cooperation. In particular, there is a cutoff for the discount factor such that if players are patient enough, then there is universal cooperation (all players cooperate with all other players); otherwise, there is no cooperation at all (Proposition 1). Conversely, when favors are substitutable, there is an upper bound on the number of relationships a player can sustain in equilibrium (Theorem 1), which is reached by most of the players if they can propose new relationships (Theorem 2), and is reached by all players who can make transfers (Proposition 5). Therefore, by considering substitutability of favors and relationships (and therefore the network's effect on the values of relationships), we can explain the empirically observed intermediate levels of cooperation even under bilateral enforcement.[3] These results show that models in which the networks determine the payoffs of individuals, but do not play an important role in enforcement may improve our understanding of social cooperation.

Heterogeneous Players and Favor Exchange Networks in Post-Soviet States. In Section 4, I extend the model to include heterogeneous players who have different costs, values, and discount factors. This allows me to distinguish between economic drivers of inequality, corresponding to the utility differences between players due to the difference in primitives such as values and costs, and social drivers of inequality, corresponding to the utility differences between players due to the network structure (see Jackson (2021) for a detailed description of these classifications and examples). Players with higher values and discount factors and lower costs have effectively *higher* types, as they benefit more from favor exchange. I show that players with high types have more relationships and are linked with other players who have high types. Therefore, they are better off not only because their types (economic drivers of inequality), but also because they can sustain more relationships and thus receive favors more frequently, and have more profitable relationships with better partners (social drivers of inequality). I then apply these results to the favor exchange networks in the Soviet Union with a particular focus on the transformation these networks

community enforcement plays an important role (examples include job referrals and sharing of information, see Jackson (2021) for a detailed description of these phenomena, as well as their role in driving inequality in societies). Finally, the large literature in relational contracts studies long-term relationships between firms that are mainly enforced through bilateral incentives as well as the legal system.

[3] For example, in the favor exchange networks in the Soviet Union described in Section 4.2, most individuals have a small set of others on whom they rely to obtain goods and services. Additionally, in favor exchange relationship data Banerjee, Chandrasekhar, Duflo, and Jackson (2013) collects in 72 Indian villages (with an average population above 900), most individuals have at least one such relationship, while more than 95 % of individuals have fewer than 5. This is also observed when players are firms: Malcomson (2010) describes how some firms cultivate long-term relationships with a subset of their suppliers.

have experienced following the increased inequality caused by the market reforms. I consider a setting with rich and poor players, where rich players can provide favors easily and are able to use transfers. I show that my model predicts three empirically observed phenomena: (i) shrinkage of the networks of the poor, (ii) the growth of the networks of the rich to the efficient levels, and (iii) the break-up of the relationships between the two groups. I also emphasize the crucial role substitutability plays in these predictions and demonstrate how the stratification of favor exchange networks can exacerbate the inequality in the society by characterizing the effects of economic and social factors explicitly.

Community and Legal Enforcement. Section 5 extends the model to incorporate community enforcement. I first describe the enforcement mechanisms used by two different groups of traders in the Medieval Era. Maghribi traders with collectivistic beliefs invested in sharing information and relied on multilateral punishment strategies, while Genoese traders with individualistic beliefs relied on a legal system that complements a bilateral punishment mechanism. I show that community enforcement crowds out bilateral enforcement; whenever community enforcement increases cooperation, it prevents the establishment of bilateral relationships with those outside of the community. This prediction is a direct consequence of substitutability and is aligned with the lack of business association between Maghribi traders and others described in Greif (1989). Next, I compare three different enforcement mechanisms: community enforcement, where players invest in sharing information; pure bilateral enforcement, where cooperation is enforced only by bilateral punishments; and legal enforcement, where players invest in a legal system that complements bilateral punishments. I characterize when each of them is optimal, shedding light on the adoption of different enforcement mechanisms by Genoese and Maghribi traders to sustain cooperation (Greif, 1994).

Related Literature. The literature on social cooperation builds on the contributions of Kandori (1992) and Ellison (1994) and focuses on community enforcement. The model in this paper is based on the one in Jackson, Rodriguez-Barraquer, and Tan (2012), who analyze the favor exchange game and focus on renegotiation-proof networks that are "robust to social contagion", which means that any breakdown in cooperation between two players spreads only to their mutual neighbors. They show that these two features imply that the network is a social quilt, which is a union of cliques. Ali and Miller (2013) focus on the other role of networks, information propagation, and show that efficient networks are cliques and cooperation is sustained through social contagion.

Bloch, Dutta, and Manea (2019) study a model in which agents search for partners by asking for favors from others and whenever a favor is provided, two agents form an exclusive long-term partnership and leave the network. Even though the games studied and the focus on bilateral enforcement are common between this paper and Bloch et al. (2019), the possibility and importance of multiple relationships allow me to focus on the size networks that can sustain cooperation, while their focus is the characterization of efficient exclusive relationships. Bendor and Mookherjee (1990) analyze bilateral enforcement in a model where all players play with each other every period, allowing players to deviate in all their relationships in a single period.[4] Xing (2016) studies a model where agents with heterogeneous autocorrelation in income form partnerships to share risk and show that this heterogeneity can contribute to inequality in society. Several other papers study networks and their effect on social cooperation through communication (Raub and Weesie (1990),Wolitzky (2013); Lippert and Spagnolo (2011); Balmaceda and Escobar (2017); Ali and Miller (2020); Sugaya and Wolitzky (2021)) or risk sharing and informal insurance (Karlan, Mobius, Rosenblat, and Szeidl (2009), Ambrus, Mobius, and Szeidl (2014), Bloch, Genicot, and Ray (2008)). Wolitzky (2021) presents a recent and detailed review of this literature.

The results in this paper complement the previous literature on social cooperation, which has mainly focused on two possible uses of networks. First, a network may reflect how players communicate with each other to exchange information about how others behave (*e.g.* Wolitzky (2013)) in order to enforce cooperation through multilateral punishments. Second, the structure of the network may play a role in sustaining cooperation through contagion strategies. Previous papers that study favor exchange (or priosoners' dilemma) models assume that players play with all players they are linked with every period and have additively separable payoffs (Lippert and Spagnolo, 2011) or each link is active (in the sense that players interact in that period) with fixed probability in each period (Jackson et al., 2012; Ali and Miller, 2013). These assumptions imply that the value players derive from a relationship is constant, does not depend on the rest of the network, and is completely lost after the relationship ends. My focus on a different role of the network, the role it plays in determining how and when players interact with each other, allows me to study some other interesting forces, such as the ability of a player to rely on the rest of their network to compensate for lost relationships.

Barron, Guo, and Reich (2022) analyze a favor exchange model where agents can accumulate wealth and show that richer agents do not have incentives to participate in favor exchange and leave the community, which they call the "too-big-for-their-boots" effect. My

[4]For example, they show that multilateral enforcement cannot improve upon bilateral enforcement when payoffs are separable and symmetric, which is not true when players are interacting with one of their partners in each period.

analysis complements theirs by showing that favor exchange networks become stratified when agents are heterogeneous (*e.g.*, rich and poor) and demonstrates inequality has an effect even when agents stay in the community.

Acemoglu and Wolitzky (2020) study a model in which agents specialized in the enforcement of cooperation may exert coercive punishment. In their model, the incentives of the specialized enforcers to carry out costly punishments are central and determine whether specialized enforcement or a mix of community and specialized enforcement is optimal. I model specialized enforcement through a legal system in a nonstrategic way, where punishment can be exerted upon the deviating agents by courts. Unless the legal system is very efficient, the use of bilateral enforcement always improves outcomes. My results on the comparison of different enforcement mechanisms complement theirs by focusing on different trade-offs such as the cost of forming communication networks and population size.

This paper is also spiritually related to the large literature on relational contracts (see Malcomson (2010) for a survey). In particular, the prediction that intermediate cooperation can be sustained by bilateral enforcement when favors are substitutable is reminiscent of the insider and outsider firms in Board (2011). In his model, faced with relationship-specific investments and hold-up problem, the principal divides agents to two groups (insiders and outsiders), while in my model, agents can sustain intermediate levels of cooperation due to the diminishing marginal value of relationships when favors are substitutable.

2. Model

There is a finite set $N = \{1, 2, ..., n\}$ of players, connected on a network described by an unweighted and undirected graph g, represented by the set of its links. I use ij to represent the link between i and j, so $ij \in g$ indicates that i and j are linked in the network g. $g - ij$ and $g + ij$ denote the networks obtained from g by deleting and adding the link ij, respectively. The neighbors of player i are denoted by $N_i(g) = \{j|ij \in g\}$ and the degree of player i in the network g is the number of their neighbors, $d_i(g) = |N_i(g)|$.

Time proceeds in discrete periods indexed by $t \in \{0, 1, \dots\}$ and g_t denotes the network at the beginning of period t. In any period, any player needs a favor with probability α.[5] F denotes the finite set of different favor types. Whenever a player requires a favor, the type of the required favor is randomly determined (with uniform distribution over F). In any period, player i is able to provide favor type $f \in F$ with probability p_{if}. Thus, j can provide favor f to i if i needs a favor of type f and j can provide favor f in that period.

[5] For simplicity, I assume at most one player requires a favor at a given period. This does not affect the results.

These probabilities are collected in a *favor provision matrix* M described in Table 1.

Table 1: Favor Provision Probabilities

$$M = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1|F|} \\ p_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ p_{n1} & \cdots & \cdots & p_{n|F|} \end{bmatrix}$$

In period t, if i requires a favor and some $j \in N_i(g_t)$ is able to provide it, then i and j play a normal form game $G = \langle \{i, j\}, \{A_1, A_2\}, \{u_1, u_2\} \rangle$, where i is player 1 with action space A_1 and utility function $u_1(a_1, a_2)$ and j is player 2 with action space A_2 and utility function $u_2(a_1, a_2)$. If multiple players in $N_i(g)$ can provide favor (of type f), then i plays the game with one of those players, who is determined randomly.[6] Therefore, a society can be described by $(N, \alpha, M, G, \delta)$.

At $t = 0$, players start the game with the network g_0. Each player knows the identity and the degree of the players that they are linked with at g_0.[7] In each period t, the timing is as follows:

1. Players decide whether to remove any links or not.
2. The player who requires the favor (player i) and the type of the required favor (f) is randomly determined.
3. The set of players who can provide f at period t is determined according to $\{p_{jf}\}_{j \neq i}$.
4. If multiple players in $N_i(g_t)$ can provide favor f, one of them (player j) is randomly selected (with uniform probability) to provide the favor.
5. Players i (as player 1) and j (as player 2) play the game G.

Players observe with whom they are playing, the action profile in every period they play, and $N_i(g_t)$ but not g_t. Strategies are denoted by $\sigma_i : H_{t,i} \times \{N \setminus i\} \rightarrow \Delta(\{K, R\} \times A_1 \times A_2)$, where $\sigma(h_{t,i}, j)$ is a probability distribution over i's decisions on her interaction with j at history $h_{t,i}$; whether they keep (K) or remove (R) the link with j and which (possibly mixed) action

[6] The assumption that i and j only plays if j is able to provide the favor can be interpreted as whether a player can perform a favor or not is observable by both players. This assumption makes defection observable, allows the study of stable networks, and facilitates comparison with related papers (such as Jackson et al. (2012), which this paper builds on) that do not focus on monitoring. The diminishing marginal value of relationships under substitutable favors would still be present even when a defection is not perfectly observable, but this would require studying different equilibria with punihsments on the equilibrium path instead of the stable networks.

[7] For each player, network g_0 matters only through the degree of their neighbors. Thus, their payoffs and incentives does not depend on the rest of the network and the results hold regardless of their knowledge about the rest of the network.

to play if they are playing with j at period t. I will study the following favor exchange game (G_f) described in Table 2, and its modifications.[8]

Table 2: Payoffs in the Favor Exchange Game

	C	D
A	$(v, -c)$	$(0, -\gamma)$

In G_f, player 1 asks for a favor from player 2, who either cooperates and provides the favor (C), or defects and refuses to provide the favor (D). If player 2 cooperates, player 1 gets v, the value of the favor, and player 2 gets $-c$, the cost of providing the favor. I assume $v > c > 0$, therefore doing favors is costly, but it is Pareto efficient for players to exchange favors over time. If player 2 refuses, then the favor is not provided and player 1 gets 0, while player 2 gets $-\gamma$, where $\gamma \in [0, c)$.[9] The term $-\gamma$ denotes any additional penalty that could be imposed on a deviating player, possibly by a legal system. In some settings, such as favor exchange networks between individuals, $\gamma = 0$ is more appropriate since the interaction has no legal implications, while in some other settings, such as the court system of Genoese traders, $\gamma > 0$ is more reasonable.

3. Bilateral Enforcement and Subsitutability

3.1. Bilateral Enforcement

To isolate the effect the network structure has on the value of each relationship, and abstract away from its effect on enforcement of cooperation, I first analyze bilateral enforcement of cooperation. Let $h_{t,i} \in H_{t,i}$ denote the period t history of player i and $h_{t,ij} \in H_{t,ij}$ denote the bilateral history between players i and j. Formally, $h_{t,ij}$ only includes the periods and action profiles of all past interactions between i and j.

Definition 1. *A strategy profile σ is* **measurable with respect to bilateral histories** *if for all i and j, $\sigma_i(h_{t,i}, j) = \sigma_i(h'_{t,i}, j)$ whenever $h_{t,ij} = h'_{t,ij}$.*

If σ is measurable with respect to bilateral histories, then the play of i when playing with j only depends on the past interactions between them, and not on their past interactions with other players, ruling out community enforcement. I will analyze *Bilateral Equilibrium*,

[8] This game corresponds to the stage game studied in Jackson et al. (2012) when $\gamma = 0$.

[9] Until Section 5, I treat γ as a fixed parameter. In Section 5, the society chooses γ by paying the cost $C(\gamma)$ of maintaining a legal system with expected punishment γ.

a refinement of Sequential Equilibrium that requires players' strategies to be measurable with respect to bilateral histories.[10]

Definition 2. (g, σ) *is a* **bilateral equilibrium** *if:*

1. *σ is measurable with respect to bilateral histories.*
2. *There exists a sequential equilibrium $(\sigma, \{\mu(h_{t,i})\}_{h_{t,i} \in H_{t,i}})$ with $g_0 = g$, strategies σ and beliefs $\{\mu(h_{t,i})\}_{h_{t,i} \in H_{t,i}}$.*

If all favors are performed on the equilibrium path, the complete network maximizes favor provision. However, players have an incentive to refuse to provide favors due to the immediate costs and delayed rewards they entail. The main goal of this paper is to characterize the *stable* networks (and its refinements) where cooperation can be sustained.

Definition 3. *A bilateral equilibrium (g, σ) is* **stable** *if all players play C (i.e., all favors are provided) and all the links are kept on the equilibrium path.*

A network g is *stable* if there exists σ_g such that (g, σ_g) is a stable bilateral equilibrium. In any stable network, the agents get the same payoff in all stable bilateral equilibria, which does not depend on the strategy σ_g. A stable network g is an *efficient stable network* if at g, all agents receive the highest payoff they can receive in any stable network.

3.2. Monopolistic Cooperation

I first study monopolistic cooperation as a benchmark. Consider the setting where the number of agents is the same as the number of favors ($|N| = |F|$) and let $M = pI \equiv M_m(p)$, where I is the identity matrix and $p \in (0, 1)$. Under $M_m(p)$, player i can only provide favor k if and only if $i = k$, in other words, each player has a monopoly over their type of favor. I refer to $M_m(p)$ as *monopolistic cooperation.* When cooperation is monopolistic, two linked players i and j interact if i (j) requires favor j (i), which happens with probability α/N, and j (i) is able to provide the favor that period, which happens with probability p.

Remark 1. *Under monopolistic cooperation, the probabaility that i and j interact at any period is independent of the rest the network and equal to $\hat{p} = \frac{\alpha p}{N}$.*

Thus, any model that determines which two linked players play in each period randomly from the uniform distribution, or assumes each link in the network is active with some exogenously given probability that does not depend on the network is equivalent to monopolistic

[10] A sequential equilibrium is a strategy profile and belief system in which, for every player i and private history $h_{t,i}$, player i's continuation strategy is optimal given her belief μ about the vector of private histories $(h_{t,j})_{j \in N}$ and there exists a sequence σ^k with $\lim_{k \to \infty} \sigma^k = \sigma$ such that for each k, μ^k is derived from σ^k using Bayes rule and $\lim_{k \to \infty} \mu_k = \mu$.

cooperation in my setting.[11] Moreover, as each additional link creates more opportunities to play at the exact same rate regardless of the network structure, the expected per period utility of player i under monopolistic cooperation, $u_i^m(g)$, is linear in $d_i(g)$:

$$u_i^m(g) = d_i(g)\hat{p}(v - c) \tag{1}$$

The following result helps explain why the earlier literature concentrated on multilateral enforcement rather than bilateral enforcement in monopolistic models.

Proposition 1. *Under monopolistic cooperation:*

- *If* $c - \gamma > \dfrac{\delta}{1-\delta}\hat{p}(v - c)$*, then the empty network is the unique stable network.*
- *If* $c - \gamma \leq \dfrac{\delta}{1-\delta}\hat{p}(v - c)$*, then the complete network is stable.*

The LHS of the inequality corresponds to the immediate cost savings a player can obtain by refusing to provide the favor, while RHS corresponds to the expected present value of future cooperation with any given player. Proposition 1 shows that in terms of networks that can be supported by bilateral enforcement, there is a sharp and unrealistic discontinuity under monopolistic cooperation: either the empty network is the unique stable network; hence, any cooperation is impossible, or the complete network is stable, and thus universal cooperation can be sustained in equilibrium. Which state prevails is determined by comparing the immediate cost of the favor with the expected present value of the relationship. Moreover, as under monopolistic cooperation the value of relationships does not depend on the rest of the network, we have the following corollary.

Corollary 1. *If* $ij \notin g$*, then* $u_i^m(g + ij) > u_i^m(g)$ *and* $u_j^m(g + ij) > u_j^m(g)$*.*

Therefore, the complete network is the unique network where no two players can form a new and mutually beneficial relationship. In Section 3.5, I formally define strong stability to characterize networks that are robust to creation of links mutually beneficial links. Corollary 1 implies that the complete network is the unique strongly stable network under monopolistic cooperation. As a result, if one does not consider substitutability of favors, bilateral enforcement makes unrealistic predictions of no cooperation or universal cooperation. In the next section, I show how incorporating substitutability enriches the model and results in intermediate level of cooperation in equilibrium.

[11] Jackson et al. (2012) assumes that in each period, any two linked players can interact with an exogenously given probability, Lippert and Spagnolo (2011) assumes that all players play with each other in all periods, while Ali and Miller (2013) assumes that each link is active with an exogenously Poisson rate. These models correspond to monopolistic cooperation in the setting of this paper. To see the equivalence formally, note that given N and the exogenous rate each link is active (r), we can compute α and p such that $\hat{p} = \frac{\alpha p}{N} = r$.

3.3. *Substitutable Favors*

This section studies substitutable favors, where all players are homogenous in their favor provision abilities. Appendix C.1 considers more general substitutable favor provision matrices and extends the main result.

Definition 4. *The favors are substitutable if* $p_{ij} = p$ *for all* i, j *and* $p \in (0, 1)$.

When favors are substitutable, assuming all favors are performed, players have the following expected utility for a period under network g:

$$u_i(g) = \alpha v \underbrace{\left(1 - \overbrace{(1-p)^{d_i(g)}}^{\substack{\text{Probability that} \\ \text{no one is available}}}\right)}_{\substack{\text{Probability that} \\ \text{favor is performed}}} - \sum_{j \in N_i(g)} \alpha c \underbrace{\left(\frac{1-(1-p)^{d_j(g)}}{d_j(g)}\right)}_{\substack{\text{Probability that favor} \\ \text{is performed by } i \text{ to } j}} \tag{2}$$

The first term represents the benefit of having $d_i(g)$ neighbors: i will need a favor worth v with probability α and there will be at least one player who can provide the favor with probability $1-(1-p)^{d_i(g)}$. The second term is the cost of having neighbors with degrees $d_j(g)$. Any $j \in N_i(g)$ receives a favor with probability $\alpha(1-(1-p)^{d_j(g)})$. As the player performing the favor is selected randomly and all players are homogeneous in their favor provision probabilities, there is a $\alpha \frac{1}{d_j(g)} \left(1-(1-p)^{d_j(g)}\right)$ chance that player i will provide it, costing her c. Two features of u_i are direct consequences of substitutability and instrumental for the analysis.

Proposition 2. *The marginal value of relationships is decreasing in the number of relationships and players with higher degree are better partners. That is,*

1. *Suppose that* $d_j(g) = d_k(g)$, $ij \notin g$ *and* $ik \notin g$. *Then*

$$u_i(g+ij) - u_i(g) > u_i(g+ij+ik) - u_i(g+ij)$$

2. *Suppose that* $ij \in g$ *and* $jk \notin g$. *Then* $u_i(g+jk) > u_i(g)$.

Proposition 2 demonstrates that the network structure is an important determinant of the value each player derives from their relationships. The marginal value of new relationships is decreasing because a relationship affects the provision of a favor when it is pivotal, *i.e.* when others are unable to provide it, and a relationship is pivotal less frequently when a player has more links. This feature is fundamentally different (and, in most cases, more realistic) than monopolistic cooperation, where the marginal value of a new relationship is constant and does not depend on how many links the players have. Moreover, players prefer their neighbors to have more links as when a neighbor is well-connected, the number of other players who can provide a favor to her is higher, which reduces the probability that any given player provides it. Thus, when favors are substitutable, relationships are more valuable for a player when they are scarce and when the partner is well connected. These channels will be important for the stratification of networks with heterogeneous players.

3.4. Stable Networks

To characterize the stable networks, I first define sustainable relationships, where the cost of providing the favor today and keeping the link is preferred by both players in a relationship to not providing the favor but losing their relationship.

Definition 5. *A relationship ij is* ***sustainable*** *at g if for $x \in \{i, j\}$,*

$$\frac{\delta}{1-\delta}u_x(g) - \frac{\delta}{1-\delta}u_x(g-ij) \geq c - \gamma \tag{3}$$

As the frequency of interactions between i and j depend on g, so does the value of the relationship. The next proposition shows that once the network's effect on the payoffs of players are taken into account, the stability of a network can be determined by evaluating the sustainability of each relationship in isolation.

Proposition 3. *A network g is stable if and only if all $ij \in g$ are sustainable at g.*

Proposition 3 shows that there exists a bilateral equilibrium (g, σ_g) such that all favors are provided and all links are kept on the equilibrium path if and only if Equation 3 is satisfied for all relationships in g. This is an important simplification in the analysis of stable networks, as instead of searching for a strategy and belief pair that constitutes a stable bilateral equilibrium under a given network g, we only need to check whether or not inequality in Equation 3 is satisfied for all relationships in g.

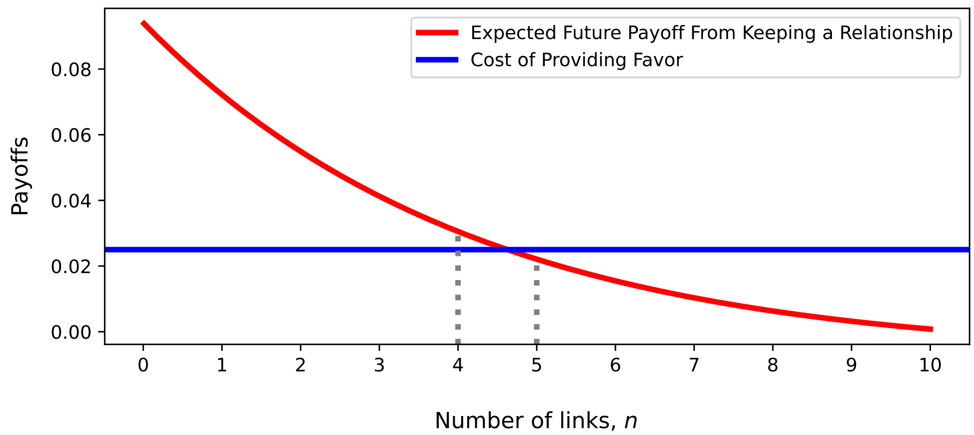
The LHS of Equation 3 is the marginal (expected, discounted) value of a relationship at g, which is decreasing when the favors are substitutable, while the RHS of 3 is the immediate cost of providing a favor (net of punishment associated with not providing the favor), which is constant. As a result, players provide favors to each other only if their degree is low enough, which limits the extent of cooperation in any stable network and results in the following theorem.

Theorem 1. *There is a $B(\alpha, p, v, c, \delta) \equiv B^*$ such that if there exists j with $d_j(g) > B^*$ then g is not stable.*

As no player can have more than B^* relationships on any stable network, I refer to B^* as the cooperation bound.[12] To gain intuition for Theorem 1, first note that as players prefer

[12] In Section C.1 in the Appendix C, I relax the more general substitutable favor provision matrices and derive a similar cooperation bound.

Fig. 1. Cooperation Bound B^*



Note: Red curve plots LHS of Equation 4 (present value of future payoff from a relationship), while the blue line plots the RHS (cost of providing a favor today). Curves intersect between 4 and 5 (dashed lines). Thus, $B^* = 4$. Parameter values: $\delta = 0.95$, $\alpha = 0.1$, $p = 0.2$, $v = 5.3$, $c = 1.5$, $\gamma = 1$.

neighbors to have more links, a player can support the maximum number of neighbors if all those neighbors have the maximum number of neighbors themselves. The following equation represents the main trade-off for a player who has n neighbors who all have n neighbors:

$$\underbrace{\delta\alpha v(1-p)^{n-1}p}_{\substack{\text{Benefit from being}\\ \text{helped in future}}} - \underbrace{\delta\alpha c\frac{1-(1-p)^n}{n}}_{\substack{\text{Cost of}\\ \text{helping in future}}} \geq \underbrace{(1-\delta)(c-\gamma)}_{\substack{\text{(Net) Cost of}\\ \text{helping today}}} \tag{4}$$

The LHS of Equation 4 is strictly decreasing in n and converges to 0, thus there exists a B^* such that the inequality in Equation 4 holds for all $n \leq B^*$ and not for any $n > B^*$. Figure 1 plots the boths sides of this inequality and illustrates how B^* is determined.

To see why no player can sustain more than B^* links, observe that if there is a player with more than B^* links in a network g, then the player who has the most links will not provide a favor if asked, and g is not stable.[13]

Although stability characterizes all possible networks that can be sustained when reached, it does not take a position on link formation. In particular, many networks in which all players have B^* or fewer links are stable, including the empty network, as there are no relationships to end or favors to perform.[14] One approach in the literature is to concentrate

[13]The reason for this is intuitive: let i denote the player who has most links in g. Then i would prefer not to provide a favor to any of her neighbors, since $d_i(g) > B^*$ and all neighbors of i have at most $d_i(g)$ links.

[14]A network where all players have fewer than B^* links can still violate stability if there are two linked players with a large difference in number of relationships. This is because the high degree player may not want to keep its relationship with a lower degree player as that player requires more frequent favor provision. For example, under the parameters in Example 1, $B^* = 9$, but any network where two linked agents who

on the most efficient networks that can be sustained in equilibrium. Equation 4 shows that the per period payoff of players is increasing in the number of relationships until they reach B^*. As a result, whenever $N \geq B^*$, which I will assume for the rest of the paper, we have the following corollary.

Corollary 2. *A network g is an efficient stable network if and only if $d_i(g) = B^*$ for all i.*

Therefore, if favors are substitutable, we expect to see intermediate levels of cooperation in equilibrium even in the absence of multilateral enforcement and any kind of monitoring technology. The next section considers a refinement of stability that allows players to form mutually beneficial relationships and shows that most players reach this bound.

Before moving refinements of stability, I study how B^* depends on the primitives of the model, given its importance of the value B^* for the results. As expected, cooperation becomes easier if the value of the favor is higher, the cost of the favor is lower, and the players are more patient.

Proposition 4. *$B(\alpha, p, v, c, \delta)$ is increasing in δ, v, and γ, and decreasing in c.*

3.5. *Strongly Stable Networks*

Strong stability requires the formation of any sustainable and mutually profitable links. When it is easy to identify and form mutually beneficial relationships, it is reasonable to expect a strongly stable network to emerge in equilibrium.

Definition 6. *A network g' is obtainable from g via deviations by $\{i, j\}$ if the following hold*

(i) $kl \in g'$ *and* $kl \notin g \implies \{k, l\} = \{i, j\}$

(ii) $kl \in g$ *and* $kl \notin g' \implies \{k, l\} \cap \{i, j\} \neq \emptyset$

Definition 6 characterizes all networks that i and j can create by forming a link between themselves and potentially removing their links with other players.[15] Two players i and j *violate strong stability* if they can form a mutually beneficial and sustainable link.

Definition 7. *Two players i, j violate strong stability at g via g' if g' is obtainable from g via deviations by $\{i, j\}$, $ij \in g'$ and $ij \notin g$ and the following hold:*

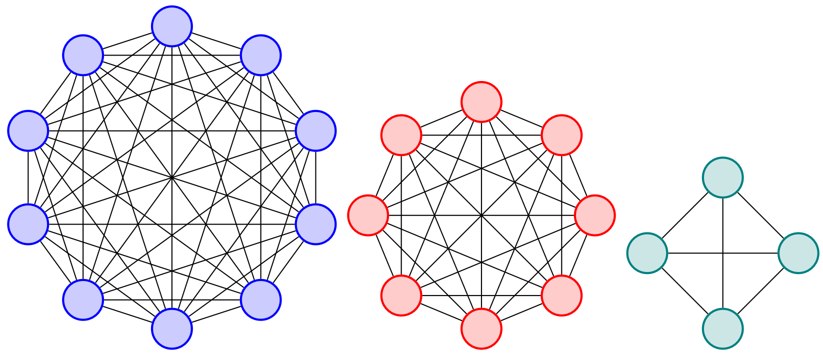
(i) ij is sustainable at g'

(ii) $u_i(g') \geq u_i(g)$, $u_j(g') \geq u_j(g)$, and at least one of the inequalities is strict.

have 8 and 5 relationships is not stable.

[15]Definition 6 corresponds to a two player restriction of the strong stability definition of Jackson and Van den Nouweland (2005). Although I consider only two player deviations, this is different from the pairwise stability introduced in Jackson and Wolinsky (1996) as their definition only checks formation and severance of links independently, not jointly.

Fig. 2. Strongly Stable Network in Example 1



A *stable* network g is *strongly stable* if there does not exist $\{i, j\}$ that violates strong stability at g.[16] The following theorem characterizes the strongly stable networks and shows that when players can form mutually beneficial and sustainable relationships, there is a bound on the number of players who do not reach B^* relationships.

Theorem 2. *For B^* in Theorem 1, the following statements are true:*

- *If $d_i(g) = B^*$ for all i, then g is strongly stable.*
- *In any strongly stable network, for any positive integer $k < B^*$, there can be at most $k + 1$ players with k links.*

The first part shows that efficient stable networks are strongly stable. Therefore, they are within the predictions of the model when profitable relationships can be identified and formed. However, there may still be some players who have fewer relationships than the efficient level. These players who do not reach the bound are in some sense undesirable partners; they require favors more frequently than they provide them. The following example illustrates this mechanism:

Example 1. *Let $p = 0.25$, $v = 7$, $c = 1$, $\alpha = 0.15$, $\gamma = 0$ and $\delta = 0.99$. It is useful to rewrite Equation 3 to denote whether a link between i (with $m = d_i(g)$) and j (with $n = d_j(g)$) is sustainable or not:*

$$h(m, n) = -c + \frac{\alpha\delta}{1-\delta}\left(v\left((1-p)^m - (1-p)^{m+1}\right) - \frac{c}{n+1}\left(1-(1-p)^{n+1}\right)\right) \quad (5)$$

A link between i and j is sustainable at g if $h(d_i(g), d_j(g)) \geq 0$. Under the given parameters, $B^* = 9$, while and $h(8, 4) < 0$. Therefore, the network in Figure 2, where three groups of players denoted by blue, red, and green nodes form three completely connected components of different sizes, is strongly stable. In this network, each blue player has 9 relationships. As $B^* = 9$ and all blue players are linked to only other blue players, no blue player can be

[16] Condition (i) is necessary since two players with same number of links can always propose an unsustainable link that makes both of them better off.

part of a coalition that violates strong stability. Each red player has 7 links and each green player has 3 links. The only possible violation of strong stability may come from a proposed link between a red and a green player. However, as $h(8,4) < 0$, any links between these two groups would not be sustainable, and the network is strongly stable.

This example illustrates how networks can exacerbate inequality, even in symmetric models. Having a low degree makes players worse off not only because they are receiving favors less frequently but also because it makes it harder for them to form new, sustainable relationships. Moreover, they have relationships with others who have low degrees, which further reduces their payoffs.

The second part of the proposition limits the number of players who do not have B^* links. To see the intuition behind this result, note that if i and j have $n < B^*$ links, $g + ij$ is sustainable and makes both players better off. Thus, they must be linked in any strongly stable network. As B^* does not depend on N, we obtain the following corollary, which shows that in large societies, almost all players reach the cooperation bound B^* when mutually profitable relationships are formed.

Corollary 3. *As $N \to \infty$, in strongly stable networks, the fraction of players who do not attain the cooperation bound vanishes.*

Theorems 1 and 2 show that when favors are substitutable, values of relationships can depend on the network structure, bilateral enforcement makes reasonable predictions of intermediate cooperation instead of full or no cooperation, and offers an alternative model to understand social cooperation.

3.6. Transfers

In many settings, players can repay each other in ways other than performing the required favors. For example, certain players may have access to funds that they can use to supplement their connections, while some others are cash constrained and could only repay through performing a future favor. Similarly, if the players represent firms and favors represent a service provided by the firms, parties can compensate each other through transfers.[17] Thus, by incorporating transfers, we can obtain insights regarding relational contracts between firms, where bilateral punishments play an important role in sustaining cooperation.[18]

[17]In such settings, the quality of the provided service may be imperfectly observable or noncontractible, so firms can still deviate by providing a subpar service, a low quality product, or refusing to share the profits, which will correspond to not providing favor.

[18]Empirical work has shown that relational contracts enforced by bilateral punishments play an important role in sustaining cooperation. McMillan and Woodruff (1999) interviews firms in Vietnam and report that

To understand the effect transfers have on cooperation and stable networks, I modify the stage game and allow player 1 (i) to either cooperate (choose action C) and pay transfer t_{ij} to player 2 (j), or defect (choose action D) and refuse to pay, where $t_{ij} \geq 0$ is the expected payment i should make to j for the favor. t denotes the matrix with ij'th entry equal to t_{ij}, which I refer as the *transfer scheme*. To model the possibility that some players aren't able to make transfers, perhaps due to budget constraints or inequality, I use $\tilde{N} \subseteq N$ to denote the set of players who can make transfers, where $t_{ij} = 0$ if $i \notin \tilde{N}$.

Table 3: Payoffs in the Favor Exchange Game with Transfers

	C	D
C	$(v - t_{ij}, t_{ij} - c)$	$(-t_{ij}, t_{ij} - \gamma)$
D	$(v - \gamma, -c)$	$(-\gamma, -\gamma)$

As transfers affect the payoffs, I refer (g, σ, t) as an equilibrium, but the definitions of bilateral equilibrium and stable networks are the same.[19] $\hat{u}_i(g, t)$ denotes the per period expected utility of the players at network g.

$$\hat{u}_i(g,t) = \alpha \left(1 - (1-p)^{d_i(g)}\right) \left(v - \sum_{j \in d_i(g)} \frac{t_{ij}}{d_i(g)}\right) - \alpha \sum_{j \in N_i(g)} \left(\frac{1 - (1-p)^{d_j(g)}}{d_j(g)}\right) (c - t_{ji}) \quad (6)$$

The definition of sustainable relationships uses $\hat{u}_i(g, t)$ instead of $u_i(g)$ and considers the most profitable profitable deviation ($\max\{c, t_{ij}\}$) for each player.

Definition 8. *A relationship ij is sustainable at g, t if for $x \in \{i, j\}$*

$$\frac{\delta}{1-\delta}\hat{u}_x(g', t) - \frac{\delta}{1-\delta}\hat{u}_x(g' - xy, t) \geq \max\{c, t_{xy}\} - \gamma \quad (7)$$

Proposition 3 extends immediately to this setting.

Corollary 4. *A network and transfer scheme g, t is stable if and only if all $ij \in g$ are sustainable at g, t.*

To extend the definition of strong stability, I allow i and j to propose bilateral transfers (t_{ij}, t_{ji}) in addition to forming a link between themselves and possibly severing their ties with other friends.

even though relational contracts play an important role in sustaining cooperation, only 19 % of the firms would expect third-party sanctions if they are deviated by one of their partners. Hendley and Murrell (2003) find that firms relied on bilateral mechanisms 56% of the time, while enforcement that relied on third parties accounted for less than 11% in a study of 254 Romanian firms. Murrell (2003) analyzes under what conditions bilateral enforcement is preferred and finds that bilateral enforcement is used when the frequency of interactions is higher and the quality of the product is important but unverifiable.

[19] For the latter, all players playing C implies that all favors are provided and all transfers are made on the equilibrium path.

Definition 9. *Two players i and j violate strong stability with transfers at g, t via g', t' if g' is obtainable from g via deviations by $\{i, j\}$, $ij \in g'$ and $ij \notin g$ and the following hold:*

(i) ij is sustainable at (g', t')

(ii) $\hat{u}_i(g', t') - u_i(g, t) \geq 0$, $\hat{u}_j(g', t') - u_j(g, t) \geq 0$, with at least one inequality strict.

(iii) $t_{xy} = t'_{xy}$ for all $\{x, y\} \neq \{i, j\}$

A network is *strongly stable with transfers* if there does not exist i and j who violates strong stability with transfers. The following proposition shows that transfers facilitate cooperation: players who can make transfers reach the cooperation bound.

Proposition 5. *Suppose that $|\tilde{N}| \geq B^*$. In a strongly stable network with transfers, all players in $\tilde{N}$ have B^* links.*

The reason behind this proposition can be explained as follows. When a player can make transfers, she can compensate others for the favors they provide. This allows the player to avoid situations where new relationships are difficult to form due to a lack of initial connections. When all players can make transfers, in any strongly stable network, all players attain their cooperation bound.

Corollary 5. *Suppose that $\tilde{N} = N$. A network is strongly stable with transfers if and only if all players have B^* links.*

Proposition 5 and Corollary 5 show that the prediction of bounded cooperation is robust to introduction of transfers and is applicable to the settings where players are firms that conduct frequent business deals. If partners (*e.g.*, suppliers) are interchangeable, then a firm with a given size can only sustain long term relationships with a bounded number of partners. This offers an alternative explanation for some firms choosing to conduct business with a limited number of long-run partners, in addition to more pronounced reasons such as the costly formation of relationships and relation-specific investments (Asanuma, 1989).

4. Heterogeneous Players and Social Networks in Post-Soviet States

This section introduces heterogeneity to the model. Section 4.1 extends the model to allow for players with different values, costs and discount factors. Players with higher values and discount factors and lower costs have *higher* types. Using the terminology presented in Jackson (2021), I differentiate between economic and social drivers of inequality. Strongly stable networks exhibit homophily and exacarbate inequality, where players with lower types are worse off not only because of economic drivers (higher costs and lower values), but also because of social drivers (they can support fewer relationships and are linked to other low type players). Section 4.2 uses these results to explain the stratification of the social networks in the post-Soviet states and demonstrate how social drivers contribute to the inequality.

4.1. Heterogeneity in Values, Costs and Discount Factor

Player i has cost c_i from performing a favor, value v_i from receving the favor and discounts future payoffs by δ_i. The vector $\nu_i = (v_i, c_i, \delta_i)$ is the type of player i.[20] Formally, the players play the following favor exchange game (where i is Player 1 and j is Player 2):

Table 4: Payoffs in the Favor Exchange Game with Heterogeneous Players

	C	D
A	$(v_i, -c_j)$	$(0, -\gamma)$

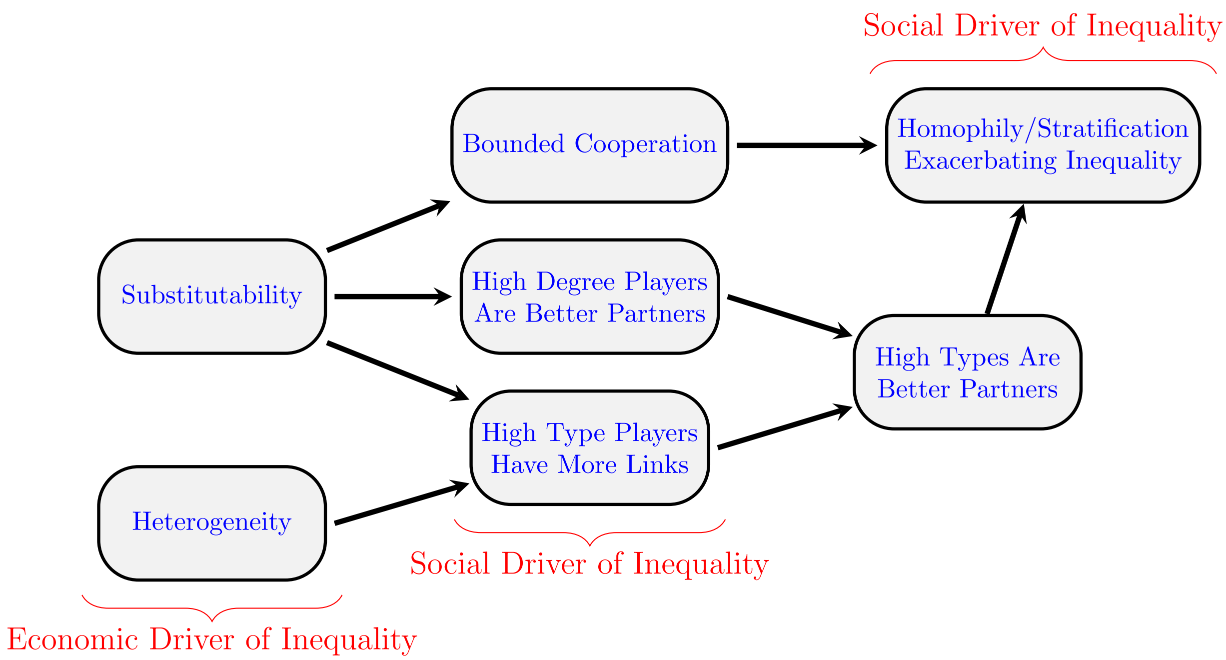
For any i, let $B(\nu_i) \equiv B(\alpha, p, v_i, c_i, \delta_i)$ denote the cooperation bound in a society formed by players of type $\nu_i \in \mathcal{V}$, where $\mathcal{V}$ is finite. We say that i's type is lower than j's type if $B(\nu_i) < B(\nu_j)$ and i's type and j's type are equivalent if $B(\nu_i) = B(\nu_j)$. The following proposition shows that most players only have links with other players of equivalent types.

Proposition 6. *In strongly stable networks, the fraction of players who attain their cooperation bound and the fraction of players who are only linked with players of equivalent types goes to 1 as $N \to \infty$.*

I illustrate the logic behind Proposition 6 and how the network structure exacerbates inequality in Figure 3. First, substitutability causes high degree players to be better partners, as they require favors less frequently. Second, when players are heterogeneous and favors are substitutable, high type players can sustain more links. This is the first channel through which the network exacerbates inequality; high type players receive favors more frequently as they have more relationships. Moreover, these two effects combine to make high type players more preferred partners in equilibrium. Third, substitutability of favors leads to the cooperation to be bounded, which combines with the fact that high types are more preferred partners and causes strongly stable networks to exhibit homophily. Therefore, low type players are not only worse off due to their higher costs or lower values and smaller network, they also have lower quality partners. In the next section, I apply this result to the favor exchange networks in Post-Soviet States and explicitly characterize all three channels.

[20] Appendix C.3 considers the alternative case where the value of the favor is determined by the player who provides the favor and extends the results.

Fig. 3. Economic and Social Drivers of Inequality



4.2. Application: Favor Exchange Networks in Post-Soviet States

I first give a short historical background on the favor exchange networks in the Soviet Union, focusing on the bilateral nature of the enforcement and the transformation these networks experienced after the dissolution of the Soviet Union. Second, I apply my results to this setting and demonstrate how the inequality caused following the dissolution of the union led to stratification of social ties and homophily in the favor exchange networks, which exacerbated the inequality.

Favor Exchange (Blat) Networks in Soviet Union. *Blat* is defined as the use of personal networks and informal contacts to obtain goods and services in short supply. It is a reciprocal relationship in which people exchange favors over time and was prevalent in the Soviet Union, where perpetual conditions of shortage and lack of access to diverse goods and services through the markets necessitated the use of such relationships.[21] These favor exchange networks were necessary to acquire many services ranging from basic necessities and quality healthcare to leisure activities such as traveling or attending concerts.[22] Ledeneva

[21]A similar institution, *Guanxi*, is present in China. As in blat relationships, Guanxi is characterized by an informal and personal connection between two individuals who adhere to an implicit psychological contract to maintain a long-term relationship based on interactions following dynamic reciprocity and long-term equity principles (see Chen and Chen (2004) for more details on Guanxi, as well as the bilateral nature of the relationships).

[22]For example, the weekly goods and services a Soviet lawyer needs to obtain by blat include "food, dry cleaning, toilet paper, concert tickets and flowers" (Simis, 1982).

(1998), based on 56 interviews of individuals involved in these relationships, describes this phenomenon as

> Blat was oriented to different needs in different historical periods (it was already flourishing in the 1930's) ... "Blat was simply a necessity for a decent life. You couldn't eat or wear what you bought in the shops, everything was in short supply, queuing and bad quality of services were appalling. To live normally, one had to have acquaintances and informal access to every sphere where needs arose" many respondents remembered.

Therefore, many individuals used their position in society to be useful to others who could reciprocate. An important aspect of these relationships is that the main enforcement mechanism was bilateral. The continuation of cooperation is primarily enforced by the threat of denial of future favors. The relationship continues as long as both parties benefit from it and is enforced by the threat of losing future exchanges.[23] These observations are summarized in Ledeneva (1998) as:

> Because blat tends to be repetitive and often operated with known partners and because of the absence of any sanctions outside the relationship, it is possible to speak of balance in blat relations.

Favor Exchange Networks in Post-Soviet Countries. The favor exchange networks have experienced changes in the market economy that followed the dissolution of the Soviet Union. For example, market reforms led to richer individuals cutting ties with poorer ones in post-Soviet Russia and Jordan.[24] Kuehnast and Dudwick (2004) present a detailed study of favor exchange networks in the Kyrgyz Republic.[25] Interestingly, the authors argue that

[23]Many interviewees who used blat emphasized this nature of the enforcement. An interviewee says: *"The exchange is fully dependent on the interest which each side has in the 'other' and previous exchanges"*. Another interviewee claims: *"... request had to be made in such a way that it would be fulfilled or, in the case of refusal, would not jeopardize the whole relationship"*, and another describes the possibility of punishment as: *"... the request is not adequate, one runs the risk of being refused or even losing the relationship"*.

[24]Two respondents in Ledeneva (1998) describe the effect of the increased inequality as *"People separate when there is a material barrier. All my friends are now businessmen, and they turned their backs on me. Not at once, they gradually distanced themselves."* and *"Those who have became wealthy have dropped out of my circle."*. Similarly, El-Said and Harrigan (2009) note how abrupt market reforms and increased poverty in Jordan caused poorer people to withdraw from social networks, contributing to social and economic polarization.

[25]Their description of these relationships are in line with Ledeneva (1998): "... informal social networks were the most important mechanisms to get things done, obtaining access to "deficit" goods and services, acquiring accurate information about events and opportunities, circumventing regulations and, in combination with bribes, gaining access to elite education, quality health care, and positions of power. This network-based economy of reciprocal favors... was an important feature of the centralized socialist economy that helped people compensate for failures of the state. ... the relatively egalitarian conditions of Soviet society enabled most people to establish far-reaching networks.

the egalitarian conditions played a role in the formation of relationships and the size of the networks.[26] However, the dissolution of the Soviet Union and the increased inequality that came with it had important effects on the structure of these social networks. Although social networks continue to be an integral part of everyday life in post-socialist Kyrgyz society, Kuehnast and Dudwick (2004) emphasize some important changes these networks have experienced:

1. The size of networks and frequency of social encounters have significantly decreased among the poor, leading to greater economic, geographic, and social isolation. Simultaneously, the non-poor have become more reluctant to provide support to poor relatives.
2. Connections are still the main currency for gaining access to public services, jobs, and higher education. However, the non-poor are able to use cash to supplement or even substitute for connections.
3. There is increasing differentiation in the form and function of social networks of the poor and the non-poor. Polarization of these networks reflects increasing socioeconomic stratification of the population.

Heterogeneous Players and Stratification of Networks. Motivated by these observations, I consider a case where the rich and the poor have different costs to provide a favor, c_r for the rich and c_p for the poor. N_r denotes the set of rich players. The rich have more resources at their disposal. They can use transfers ($\tilde{N} = N_r$) and can grant favors easier: $c_r < c_p$. In Appendix C.3 I consider the case where rich players can provide higher value favors in a setting where the value of the favor depends on the identity of provider and extend Proposition 7.

$B^*(c) \equiv B(\alpha, p, v, c, \delta)$ denotes the cooperation bound of a society with cost c. Given c_p, let $\tilde{c}_r = \inf\{c_r \in [0, c_p] : B^*(c_r) = B^*(c_p)\}$ denote the minimum cost level under which rich and poor players can support the same number of relationships.[27] If $c_r > \tilde{c}_r$, then inequality is low and the rich and poor players have equivalent types, whereas if $c_r < \tilde{c}_r$, inequality is high and poor players have lower types than rich players. The next result follows from Proposition 6 and shows that the relationships between rich and poor players will be limited in strongly stable networks when inequality is high.

Corollary 6. *The following statements are true:*

- *If inequality is low ($c_r > \tilde{c}_r$), then all networks where all players have $B^*(c_r)$ links are strongly stable.*

[26] Indeed, in my model, when players are homogeneous, any two players can have a favor exchange relationship in a stable network.

[27] $\tilde{c}_r$ is characterized in the proof of Proposition 7.

- *If inequality is high ($c_r < \tilde{c}_r$) in any strongly stable network, the fraction of rich players who have $B^*(c_r)$ links, the fraction of poor players who have $B^*(c_p)$ links and the fraction of players who only have links with their own group goes to 1 as $N \to \infty$.*

Moreover, when the rich can make transfers, they can use them to make sure they reach the cooperation bound and only have relationships with other rich players. Therefore, even though transfers facilitate cooperation in symmetric models by helping formation of new links, they may also play a role in increasing the stratification of social networks.

Proposition 7. *The following statements are true:*

- *If $c_r > \tilde{c}_r$, then all networks where all players have B^* links are strongly stable with transfers.*
- *If $c_r < \tilde{c}_r$ and g is strongly stable with transfers, at g, all rich players have $B^*(c_r)$ links and none are linked with a poor player, who have at most $B^*(c_p)$ links.*

When $c_r \geq \tilde{c}_r$, the inequality is low that both types of players can sustain the same number of relationships, and heterogeneity does not have an effect on the network structure. However, when $c_r < \tilde{c}_r$, the inequality is high and lower relative costs allow rich players to sustain a larger network composed of other rich players. To illustrate clearly the three main channels that make the poor worse off compared to the rich, suppose that the costs for rich allow them to support one additional relationship: $B^*(c_p) = B$ and $B^*(c_r) = B + 1$. The per period payoff of a poor player who reaches the cooperation bound B is

$$u_p = \alpha v(1 - (1-p)^B) - \alpha c_p \left(\frac{1 - (1-p)^B}{B} \right) B \tag{8}$$

whereas the per period payoff of a rich player (who is certain to reach the cooperation bound $B + 1$) is

$$u_r = \alpha v(1 - (1-p)^{B+1}) - \alpha c_r \left(\frac{1 - (1-p)^{B+1}}{B+1} \right) (B+1) \tag{9}$$

The difference in the per period payoff of the poor and the rich can be decomposed to three channels as follows:

$$\begin{aligned} u_p - u_r = \quad & \underbrace{-\alpha \left(\frac{1 - (1-p)^{B+1}}{B+1} \right) B(c_p - c_r)}_{\text{Economic Driver of Inequality: Higher costs } \implies \text{ lower payoff}} \\ & \underbrace{-\alpha \left[(1-p)^B pv - \left(\frac{1 - (1-p)^{B+1}}{B+1} \right) c_r \right]}_{\text{Social Driver of Inequality 1: Smaller Network } \implies \text{ Lower probability of obtaining the favor}} \\ & \underbrace{-\alpha \left[\left(\frac{1 - (1-p)^B}{B} \right) - \left(\frac{1 - (1-p)^{B+1}}{B+1} \right) \right] Bc_p}_{\text{Social Driver of Inequality 2: Linked with other poor players } \implies \text{ Provide favors more frequently}} \end{aligned} \tag{10}$$

The first line of Equation 10 represents the first channel, which stems from the economic differences between the players: the higher cost of favors directly reduces the benefits of favor exchange for the poor players. The other two channels correspond to the social drivers of inequality. Second line characterizes the loss from the lower number of relationships that poor players can support compared to rich players, due to their higher costs. With fewer relationships, favors are performed less frequently for the poor even if we account for the favors the rich players needs to perform for their additional relationship.[28] Finally, the third line corresponds to the loss of utility for the poor players due to the identity of their links: they are linked to other poor players who have fewer relationships and depend more on them.[29] Since it is better to have players with more connections, the networks of poor players are not only smaller, but also composed of worse partners compared to rich players. Moreover, both the larger networks of rich players and their payoff gain from being linked to other rich players are direct consequences of substitutability and would not be obtained in a model with monopolistic cooperation, illustrating how inequality can result in the stratification of networks, which can further exacerbate the inequality in society when relationships are substitutable.

The predictions in Proposition 7 are in line with the main observations of Kuehnast and Dudwick (2004). During the Soviet era, the society was egalitarian and most individuals had networks where they exchanged favors, which is the equilibrium in the homogeneous case. However, following the market reforms, society became more unequal, social networks became more polarized, and the size of the networks of the poor decreased, exhibiting the changes predicted by Proposition 7. The implications of the proposition are illustrated in Figure 4 that depicts strongly stable networks with transfers under with and without heterogeneity in costs.

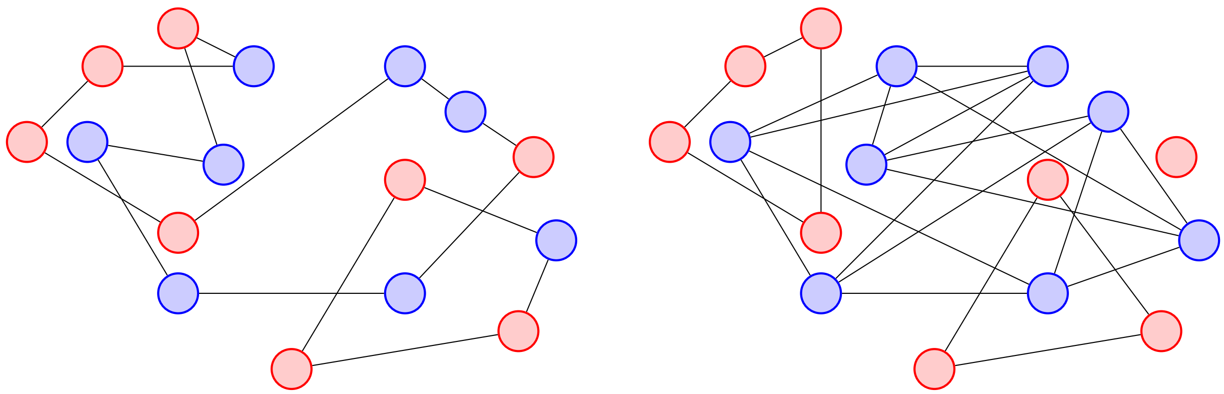
5. Community and Legal Enforcement

In this section, I first describe how and why two contemporaneous trader groups (Maghribi and Genoese traders) have adopted two different enforcement mechanisms. Second, I model community enforcement and show that when effective, community enforcement crowds out bilateral relationships, which explains the lack of relationships between foreign traders and

[28] With probability $(1-p)^B p$, a rich player receives a favor while a poor player does not. The expression in brackets in the second line corresponds to the LHS of Equation 4 and is greater than $(\delta/(1-\delta))(c-\gamma)$ as $B+1$ relationships are sustainable for rich players.

[29] The expression in brackets in the third line is positive as $\frac{1-(1-p)^n}{n}$ is a decreasing function of n for $p \in (0,1)$.

Fig. 4. Strongly Stable Networks Before and After Heterogeneity



Note: Parameter values: $\delta = 0.95$, $\alpha = 0.1$, $p = 0.1$, $v = 9$, $\gamma = 0$. LHS is a strongly stable network when all players have costs $c = 1.3$, with $B^*(c) = 2$. In this network, all players have 2 relationships, and any such network would be strongly stable. RHS is a strongly stable network where red (poor) players have costs $c_p = 1.3$, but blue (rich) players have $c_r = 1$, with $B^*(c_r) = 4$. In this network (and all other strongly stable networks with transfers), all rich players have exactly 4 links, and they are only linked to each other. Poor players lose all links with rich players but may form new links with other poor players, even though not all are guaranteed to reach 2 links in strongly stable networks.

Maghribi traders, who relied on community enforcement. Finally, I introduce costs for maintaining communication in a community and establishing a legal system, characterize the optimal mechanisms, and relate it to the choices made by Maghribi and Genoese traders.

5.1. *Maghribi and Genoese Traders*

Maghribi traders are a group of Jewish traders who lived in the Mediterranean in the eleventh century and used a reputation-based institution to deal with the contractual problems inherent in merchant-agent transactions (see Greif (2006) for a detailed description of the institutions Maghribi traders have used). Maghribi traders mainly relied on a multilateral reputation mechanism, where traders share information about each others' behavior with other traders and use a collective punishment strategy of not trading with agents who have previously cheated others. However, another group operating in the same era, the Genoese traders, have adopted a different enforcement mechanism that does not rely on collective punishments. Greif (1994) summarizes these differences as follows:

> Collectivist cultural beliefs were a focal point among the Maghribis, and individualist cultural beliefs were a focal point among the Genoese. ... The historical evidence indicates that the Maghribis invested in sharing information and the Genoese did not. Each Maghribi corresponded with many other Maghribi traders

> by sending informative letters to them with the latest available commercial information and "gossip", including whatever transpired in agency relations among other Maghribis. Important business dealings were conducted in public, and the names of the witnesses were widely publicized. Although, most likely, not every Maghribi trader was familiar with all the others, belonging to the Maghribis was easily verifiable through common acquaintances, an extensive network of communication, a common religion, and a common language.

Unlike Maghribi traders, Genoese traders have relied on a legal system and a bilateral reputation mechanism. This mechanism is described in Greif (2006) as:

> Thus for agency relations to be established in an individualistic society, an external mechanism such as a legal system backed by the state is needed to limit agents ability to embezzle merchants' capital. A legal system complements an institution based on individualistic cultural beliefs; it does not replace the associated bilateral reputation institution. Where a legal system has only a limited ability to restrict cheating (e.g., from misreporting profit expenses), a reputation mechanism still has to be used. The extensive writing of agency contracts suggests that this was indeed the case among the Genoese.

Therefore, Genoese traders enforced cooperation through the possibility of court punishments and the threat of losing the bilateral relationship.[30] Moreover, Greif (1989) emphasizes that "Evidence of business association between Maghribi traders and non-Maghribi traders (Jewish or Muslim) is rare" and that their coalition is closed to outsiders. He attributes this to the fact that relatively higher short-run gains Maghribi traders obtain from cheating a foreign trader. In the next section, I show that when cooperation is substitutable and community enforcement enhances cooperation, it crowds out the formation of bilateral relationships, explaining the lack of relationships between Maghribi traders and outsiders.

5.2. Community Enforcement is Incompatible with Bilatreal Enforcement

Let $\Phi = \{\phi_1, \ldots, \phi_n\}$ denote a partition of the players. Each $\phi \in \Phi$ denotes a *community.* I use $\phi(i)$ to denote the community of a player i. A community $\phi \in \Phi$ is a *large community* if $|\phi| > B^*$ and is a *small community* otherwise. $\overline{\Phi} \subseteq \Phi$ denotes the set of large communities. A community is interpreted as a group of players who have invested in monitoring and/or

[30] "...among the Genoese an agent was induced to be honest by the fear that, if he was not, his relations with a particular merchant (who could have represented a family or even a clan) would be terminated (Greif, 2002). "

communication so that information about the interactions between community members can be monitored and credibly communicated within the community.

Formally, $\tilde{\mathcal{H}}$ denotes the set of all *community histories* and let $\tilde{h}_{t,ij}$ denote the history between i and j. If $\phi(i) = \phi(j)$, then $\tilde{h}_{t,ij}$ includes the periods and action profiles of all previous interactions between i, j and any $k \in \phi(i)$. Otherwise, $\tilde{h}_{t,ij}$ only includes the periods and action profiles of all past interactions between i and j, in other words, it is the bilateral history between i and j.

Definitions 1, 2 and 3 are updated by replacing $h_{t,ij}$ with $\tilde{h}_{t,ij}$ and changing the word bilateral to community. I say that a network g is *community stable* if there exists (g, σ_g) that is a stable community equilibrium.

The following proposition shows that small communities cannot increase cooperation compared to bilateral enforcement and whenever community enforcement increases cooperation, it prevents the establishment of bilateral relationships.

Proposition 8. *Suppose that g is a community stable network. Then the following statements are true:*

- *If $\phi(i)$ is a small community, then $d_i(g) \leq B^*$.*
- *If $d_i(g) > B^*$, then $\phi(i)$ is a large community, and $ij \in g$ only if $j \in \phi(i)$.*

To gain intuition for this result, suppose that i and j are linked but not in the same community. Then their relationship must be enforced bilaterally. This implies that neither player can sustain more than B^* relationships, since in that case they will strictly prefer not to provide the favor and lose the relationship. Next, suppose that $d_i(g) > B^*$. Then i is already cooperating with more than B^* players, and losing a particular relationship is much less important for i compared to another player who has fewer than B^* relationships. Therefore, i will prefer not to provide a favor to any j who is not in the same community, as this could only cause the loss of a relationship with j and would not affect any other relationship. As a result, all neighbors of i at g must belong to the same community as i and community enforcement cannot complement bilateral enforcement, it can only replace it.

This result depends crucially on the diminishing value of relationships and would not be obtained under monopolistic cooperation. If cooperation is monopolistic, the marginal benefit of establishing a bilateral relationship with a player outside the community does not depend on cooperation within the community, and this relationship may be sustainable, regardless of the size of the community. This result offers an explanation for the lack of relationships between Maghribi traders and traders from other groups.

Proposition 8 also shows that communities must be large enough to have an effect on the level of cooperation. As community enforcement and bilateral enforcement cannot be effective together, small communities cannot improve cooperation. Any player whose community is not large enough needs bilaterally enforced relationships and cannot have more than B^* links. This helps explain why the Genoese traders did not rely on community enforcement to complement their bilateral reputation mechanism and legal system.

5.3. *Efficiency of Different Enforcement Mechanisms*

In this section, I will formally define and compare three different mechanisms that can be used to sustain cooperation: pure bilateral enforcement, community enforcement, and legal enforcement.

Pure Bilateral Enforcement. Pure bilateral enforcement represents cases where cooperation cannot be enforced by a legal system or multilateral punishments. As outlined in Section 4.2, many blat relationships and favor exchange relationships between individuals in modern societies fall under this category. This is the simplest enforcement mechanism in which players play the favor exchange game with $\gamma = 0$ and cooperation is sustained only by the threat of losing the relationship. The expected payoff per period under pure bilateral enforcement is given by plugging in $d_i(g) = B^*$ in Equation 2.

Legal Enforcement. In legal enforcement, players play the favor exchange game with $\gamma \geq 0$ and each player pays $C(\gamma)$ per period to maintain this legal system. C is convex, $C(\gamma) > 0$ for all γ and $\lim_{\gamma \to c} C(\gamma) = \infty$. Legal enforcement is a good model for Genoese traders and modern firms, especially when the quality of goods and services is important and not easily verifiable.

We can interpret legal enforcement in multiple ways. First, if the courts are already present, $C(\gamma)$ can represent the costs associated with making contractual agreements that could be enforced in courts.[31] $C(\gamma)$ can also represent the per-player costs associated with maintaining the legal system.[32] The efficient networks under legal enforcement is characterized in the following corollary, which follows from Proposition 4 and Corollary 5.

Corollary 7. *For each γ, there is a $B^*(\gamma)$ such that a network is an efficient stable network if and only if all players have $B^*(\gamma)$ links. $B^*(\gamma)$ is increasing in γ. Moreover, these are the only strongly stable networks with transfers.*

Let g_γ denote the network in which all players have $B^*(\gamma)$ relationships, where all players obtain a per period payoff of $u(g_\gamma) \equiv u_i(g_\gamma)$. Then the per period payoff of a player under

[31] Indeed, Greif (1994) discusses that Genoese traders used documents called "bill of landing" to keep official records of their transactions, while Maghribi traders did not as they have solved this problem through their informal enforcement mechanisms.

[32] In Appendix C.4, I modify the per-period cost to $C(\gamma)/N$, where N is the population size and show that legal enforcement is optimal in large societies.

legal enforcement, which is attained with the network g_γ, is $U(\gamma) \equiv u(g_\gamma) - C(\gamma)$. The optimal level of legal enforcement is given by $\gamma^* = \arg\max_{\gamma \in [0,c]} U(\gamma)$.

Community Enforcement. In community enforcement, each player belongs to a community and is linked to all members of their community. Players play the favor exchange game with $\gamma = 0$. Any player i pays a per period cost κ for each j with $j \in \phi(i)$, which denotes the cost of maintaining links that allow for monitoring and communication channels.

Let $U(|\phi|, \kappa)$ denote the expected payoff per period of a member of a community of size $|\phi|$, including the costs associated with maintaining community links.

Proposition 9. *There exists a $B(\kappa)$ such that $U(|\phi|, \kappa)$ is maximized at $|\phi| = B(\kappa)$. $B(\kappa)$ is decreasing in κ. There are κ_P and κ_L such that:*

- *Community enforcement with optimal community size $B(\kappa)$ results in a higher payoff than pure bilateral enforcement if and only if $\kappa \leq \kappa_P$. Moreover, whenever $\kappa \leq \kappa_P$, $B(\kappa) > B^*$.*
- *Community enforcement with optimal community size $B(\kappa)$ results in a higher payoff than optimal legal enforcement if and only if $\kappa \leq \kappa_L$.*

When cooperation is enforced by communities, the (per-period) cost of adding a new member is κ, while the marginal benefit decreases as the community grows as favors are substitutable. Therefore, it is optimal for players to organize themselves as communities of $B(\kappa)$. For a community to increase the payoff of its members, it must be larger than the number of relationships that can be sustained by pure bilateral enforcement, which happens when the cost of maintaining the community links is low enough.

The second part of the proposition echoes Greif's observation about collectivism and individualism being a determinant of these different choices made by different groups of traders. Maghribi traders share a common religion and language, belong to a collectivistic culture, lived in a smaller society (which would correspond to lower κ), and did not integrate into the larger Jewish community in which they live (possibly to keep facilitate credible communication and κ low) as long as they are active in trade. These features led to the adoption of community enforcement. On the contrary, Genoese traders lived in a larger, individualistic society, which precluded the adoption of community enforcement and paved the way for the development of a legal system.

6. Conclusion

This paper introduces a framework to study favor exchange that allows favors to be substitutable. When favors are substitutable, the network structure determines how frequently individuals interact with each other and affects the value of their relationship. To emphasize this role of the network, I first focus on bilateral enforcement where network structure does not play a role in enforcing relationships, its other main effect on cooperation. I show that assuming any two players play in each period with an exogenously given probability corresponds to monopolistic favors in my setting and causes the equilibrium networks to exhibit either universal cooperation or no cooperation. Conversely, when favors are substitutable, the equilibrium cooperation in the society is intermediate and varies continuously with the primitives, instead of jumping from no cooperation to universal cooperation.

When players are heterogeneous, higher type players with smaller costs or higher values can sustain more relationships, which make them better partners. When favors are substitutable, players prefer higher type partners, which creates homophily and stratification in the network structure. I show that low type players are not only worse off due to higher costs and lower values (representing economic factors) but also because of their smaller networks that consist of other low type players (social factors driven by the network structure). I apply the results to favor exchange networks in the Soviet Union and show that the predictions of the model match the transformation these networks experienced after market reforms: The networks of the rich grow and consist of other rich individuals, while the networks of the poor contract and consist of other poor individuals. I also demonstrate how economic and social factors create inequality by decomposing the equilibrium payoffs.

I also study different enforcement mechanisms: community enforcement in which players can invest in communication and use multilateral punishments or legal enforcement in which players can invest in a legal system that punishes deviators. I show that bilateral enforcement and community enforcement cannot be effective together: community enforcement prevents establishment of bilateral ties and crowds out bilateral enforcement. This result offers an explanation for the lack of (bilateral) relationships between Maghribi traders who practice community enforcement and outsiders. I also show that community enforcement is optimal when communication channels are easy to maintain, which is the case for a tightly knit society like the Maghribi traders, but not for the individualistic and larger society of Genoese traders, echoing the mechanisms adopted by Genoese and Maghribi traders described in Greif (1994).

References

Acemoglu, D., Wolitzky, A., 2020. Sustaining cooperation: Community enforcement versus specialized enforcement. Journal of the European Economic Association 18, 1078–1122.

Ali, S. N., Miller, D. A., 2013. Enforcing cooperation in networked societies .

Ali, S. N., Miller, D. A., 2020. Communication and cooperation in markets. arXiv preprint arXiv:2005.09839 .

Ambrus, A., Mobius, M., Szeidl, A., 2014. Consumption risk-sharing in social networks. American Economic Review 104, 149–82.

Asanuma, B., 1989. Manufacturer-supplier relationships in japan and the concept of relation-specific skill. Journal of the Japanese and international economies 3, 1–30.

Balmaceda, F., Escobar, J. F., 2017. Trust in cohesive communities. Journal of Economic Theory 170, 289–318.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., Jackson, M. O., 2013. The diffusion of microfinance. Science 341.

Barron, D., Guo, Y., Reich, B., 2022. Wealth dynamics in communities. Working Paper .

Bendor, J., Mookherjee, D., 1990. Norms, third-party sanctions, and cooperation. JL Econ & Org. 6, 33.

Bloch, F., Dutta, B., Manea, M., 2019. Efficient partnership formation in networks. Theoretical Economics 14, 779–811.

Bloch, F., Genicot, G., Ray, D., 2008. Informal insurance in social networks. Journal of Economic Theory 143, 36–58.

Board, S., 2011. Relational contracts and the value of loyalty. American Economic Review 101, 3349–67.

Chen, X.-P., Chen, C. C., 2004. On the intricacies of the chinese guanxi: A process model of guanxi development. Asia Pacific Journal of Management 21, 305–324.

Coleman, J. S., 1994. Foundations of social theory. Harvard university press.

El-Said, H., Harrigan, J., 2009. you reap what you plant: Social networks in the arab worldthe hashemite kingdom of jordan. World development 37, 1235–1249.

Ellison, G., 1994. Cooperation in the prisoner's dilemma with anonymous random matching. The Review of Economic Studies 61, 567–588.

Granovetter, M., 1985. Economic action and social structure: The problem of embeddedness. American journal of sociology 91, 481–510.

Greif, A., 1989. Reputation and coalitions in medieval trade: evidence on the maghribi traders. The journal of economic history 49, 857–882.

Greif, A., 1994. Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. Journal of political economy 102.

Greif, A., 2002. Economic history and game theory. Handbook of game theory with economic applications 3, 1989–2024.

Greif, A., 2006. Institutions and the path to the modern economy: Lessons from medieval trade. Cambridge University Press.

Hendley, K., Murrell, P., 2003. Which mechanisms support the fulfillment of sales agreements?: Asking decision-makers in firms. Economics Letters 78, 49–54.

Jackson, M., Wolinsky, A., 1996. A strategic model of economic and social networks. Journal of Economic Theory 71, 44–74.

Jackson, M. O., 2021. Inequality's economic and social roots: The role of social networks and homophily. Available at SSRN 3795626 .

Jackson, M. O., Rodriguez-Barraquer, T., Tan, X., 2012. Social capital and social quilts: Network patterns of favor exchange. American Economic Review 102, 1857–97.

Jackson, M. O., Van den Nouweland, A., 2005. Strongly stable networks. Games and Economic Behavior 51, 420–444.

Kandori, M., 1992. Social norms and community enforcement. The Review of Economic Studies 59, 63–80.

Karlan, D., Mobius, M., Rosenblat, T., Szeidl, A., 2009. Trust and social collateral. The Quarterly Journal of Economics 124, 1307–1361.

Kuehnast, K. R., Dudwick, N., 2004. Better a hundred friends than a hundred rubles?: social networks in transition–the Kyrgyz Republic, vol. 41181. World Bank Publications.

Ledeneva, A., 1998. Russia's economy of favours: Blat, networking and informal exchange, vol. 102. Cambridge University Press.

Lippert, S., Spagnolo, G., 2011. Networks of relations and word-of-mouth communication. Games and Economic Behavior 72, 202–217.

Malcomson, J., 2010. Relational incentive contracts .

McMillan, J., Woodruff, C., 1999. Dispute prevention without courts in vietnam. Journal of law, Economics, and Organization 15, 637–658.

Murrell, P., 2003. Firms facing new institutions: transactional governance in romania. Journal of Comparative Economics 31, 695–714.

Raub, W., Weesie, J., 1990. Reputation and efficiency in social interactions: An example of network effects. American journal of sociology 96, 626–654.

Simis, K. M., 1982. Ussr: the corrupt society : the secret world of soviet capitalism .

Sugaya, T., Wolitzky, A., 2021. Communication and community enforcement. Journal of Political Economy 129, 2595–2628.

Wolitzky, A., 2013. Cooperation with network monitoring. Review of Economic Studies 80, 395–427.

Wolitzky, A., 2021. Cooperation in large societies. Working Paper .

Xing, Y., 2016. Who shares risk with whom and how? endogenous matching and selection of risk sharing equilibria. Tech. rep., SIEPR Discussion Paper.

Appendices

A. Proofs

A.1. Proof of Proposition 1

To prove the first part, assume $c - \gamma \leq \dfrac{\delta \hat{p}(v-c)}{1-\delta}$. We will show the following strategy-belief pairs are a bilateral equilibrium under the complete network:

$$\sigma_i^*(h_{t,i}, j) = \begin{cases} \{K, C\} & \text{if } ij \in N_i(g_t) \text{ and all favors are performed between } i \text{ and } j \text{ at } h_{t,i} \\ \{R, D\} & \text{otherwise} \end{cases}$$

and μ_i^* puts positive probability only to the private histories of other players where all previous favors have been performed and links have been kept apart from the ones i have already observed otherwise.

Let σ_i denote another strategy that is different from σ_i^*. In what follows, we will show that, for each $h_{t,i}$ such that σ_i deviates from σ_i^* and $j \neq i$, replacing $\sigma_i(h_{t,i}, j)$ with $\sigma_i^*(h_{t,i}, j)$ weakly increases the payoff of i, thus σ_i is not a profitable deviation. Let $i \neq j$ denote another player.

Case 1: There was a previous deviation in $h_{t,ij}$: In this case, the link $\{ij\}$ cannot be active in period t as $\sigma_j^*(h_{t,j}, i) = \{R, D\}$. As a result, replacing $\sigma_i(h_{t,i}, j)$ with $\{R, D\}$ does not change the utility of player i.

Case 2: There was not a previous deviation in $h_{t,ij}$:

Let ω_R denote the probability that R is played under $\sigma_i(h_{t,i}, j)$ and let ω_D denote the probability that $\{K, D\}$ is played under $\sigma_i(h_{t,i}, j)$. Note that the difference in the expected continuation payoff between σ^* and σ is given by

$$\omega_R \left(\frac{1}{1-\delta} \frac{\alpha p}{N} (v-c) \right) + \omega_D \frac{\alpha p}{N} \left(-c + \gamma + \frac{\delta}{1-\delta} \frac{\alpha p}{N} (v-c) \right) > 0 \tag{11}$$

where the first term is positive since $v - c > 0$ and second is positive from assumption above and the inequality follows. Thus, in both cases, σ^* does better than σ and σ is not a profitable deviation, proving the first part of the result.

To prove the second part, assume $c - \gamma > \dfrac{\delta \hat{p}(v-c)}{1-\delta}$ and that there is a stable network g such that $ij \in g$. Thus, there is a bilateral equilibrium strategies σ such that when all players follow σ, all favors are provided on the equilibrium path. Clearly, σ_i assigns $\{K, C\}$ to both players in all $h_{t,ij}$ with no previous deviation, as otherwise g would not be stable. Define σ' by replacing $\sigma_i(h_{0,i}, j) = \{K, C\}$ with $\{K, D\}$ and leaving the rest of the strategy unchanged. The probability that i requires favor j is α/N, and j is able to provide that favor with probability p. Thus, by Corollary 8 in Appendix B, the difference in expected continuation payoff between σ and σ' is given by:

$$\frac{\alpha p}{N}\left(-c+\gamma+\frac{\delta}{1-\delta}\frac{\alpha p}{N}(v-c)\right)=\hat{p}\left(-c+\gamma+\frac{\delta}{1-\delta}\hat{p}(v-c)\right)<0 \tag{12}$$

where the last inequality follows from our assumption above. As a result, σ' is a profitable deviation and there cannot exits such an equilibrium.

A.2. Proof of Proposition 2

To prove the first part, let $\hat{U}=[u_i(g+ij)-u_i(g)]-[u_i(g+ij+ik)-u_i(g+ij)]$ and $x=1-p$. Note that

$$\begin{aligned}\hat{U}&=\alpha v[(x^{d_i(g)}-x^{d_i(g)+1})-(x^{d_i(g)+1}-x^{d_i(g)+2})]\\&=\alpha v x^{d_i(g)}[1-x-x+x^2]=\alpha v x^{d_i(g)}(1-x)^2>0\end{aligned} \tag{13}$$

To prove the second part, note that

$$\begin{aligned}u_i(g+jk)-u_i(g)=&\alpha c\left(\frac{1-(1-p)^{d_j(g)+1}}{d_j(g)+1}-\frac{1-(1-p)^{d_j(g)}}{d_j(g)}\right)\\&+\mathbb{I}\left\{ik\in d_i(g)\right\}\alpha c\left(\frac{1-(1-p)^{d_k(g)+1}}{d_k(g)+1}-\frac{1-(1-p)^{d_k(g)}}{d_k(g)}\right)>0\end{aligned} \tag{14}$$

where the inequality follows as $\frac{\partial}{\partial x}\frac{1-(1-p)^x}{x}=\frac{(1-p)^x(1-x\log(1-p))}{x^2}>0$ whenever $(1-p)\in(0,1)$ and $x\geq 1$.

A.3. Proof of Proposition 3

Lemma 1. *Suppose that $ij\in g\cap\tilde{g}$, $d_i(\tilde{g})\leq d_i(g)$ and $d_j(\tilde{g})=d_j(g)$. Then $u_i(\tilde{g})-u_i(\tilde{g}-ij)\geq u_i(g)-u_i(g-ij)$.*

Proof. Observe that $u_i(\tilde{g})-u_i(\tilde{g}-ij)-(u_i(g)-u_i(g-ij))=\alpha v\Big((1-p)^{d_i(\tilde{g})}-(1-p)^{d_i(g)}\Big)\geq 0$, where the equality follows from $d_j(\tilde{g})=d_j(g)$ and the inequality follows from $d_i(\tilde{g})\leq d_i(g)$ and $p\in(0,1)$. □

To prove the if part of the proposition, let g be a network such that all links are sustainable (*i.e.*, the inequality in equation 3 holds for all $ij\in g$). We will show that the following strategy-belief pairs constitute a bilateral equilibrium:

$$\sigma_i(h_{t,i},j)=\begin{cases}\{K,C\} & \text{if } ij\in N_i(g_t) \text{ and all favors are performed between } i \text{ and } j \text{ at } h_{t,i}\\ \{R,D\} & \text{otherwise}\end{cases}$$

and μ_i^* puts positive probability only to the private histories of other players where all previous favors have been performed and links have been kept apart from the ones i have already observed otherwise.

Let σ_i denote another strategy that deviates from σ_i^* at a history $h_{t,i}$. In what follows, we will show that, for each $j \neq i$, replacing $\sigma_i(h_{t,i}, j)$ with $\sigma_i^*(h_{t,i}, j)$ weakly increases the payoff of i, thus σ_i is not a profitable deviation. Let $i \neq j$ denote another player.

Case 1: There was a previous deviation in $h_{t,ij}$: In this case, the link $\{ij\}$ cannot be active in period t as $\sigma_j^*(h_{t,j}, i) = \{R, D\}$. As a result, replacing $\sigma_i(h_{t,i}, j)$ with $\{R, D\}$ does not change the utility of player i.

Case 2: There was not a previous deviation in $h_{t,ij}$:

Let $\mathcal{D}(h_{t,i})$ denote the set of all k such that there was a previous deviation in $h'_{t,ik}$, where h' is a subhistory of h. Let $\tilde{g}$ denote the network obtained by removing all links ik for $k \in \mathcal{D}(h_{t,i})$. Note that $\mu_i^*(h_{t,i})$ puts positive probability to histories where all favors are provided and all links kept between players l and l' unless $l \in \mathcal{D}(h_{t,i})$ and $l' = i$. Let ω_R denote the probability that R is played under $\sigma_i(h_{t,i}, j)$ and ω_D denote the probability that $\{K, D\}$ is played under $\sigma_i(h_{t,i}, j)$. The difference in the expected continuation payoff at history h_t between σ^* and σ is given by:

$$\omega_R \left[\frac{\delta}{1-\delta} \left(u_i(\tilde{g}) - u_i(\tilde{g} - ij) \right) \right] + \frac{\omega_D \alpha (1 - (1-p)^{d_j(g)})}{d_j(\tilde{g})} \left[-c + \gamma + \frac{\delta}{1-\delta} \left(u_i(\tilde{g}) - u_i(\tilde{g} - ij) \right) \right] \geq$$
$$\omega_R \left[\frac{\delta}{1-\delta} (u_i(g) - u_i(g - ij)) \right] + \frac{\omega_D \alpha (1 - (1-p)^{d_j(g)})}{d_j(g)} \left[-c + \gamma + \frac{\delta}{1-\delta} (u_i(g) - u_i(g - ij)) \right] \geq 0$$

where the first inequality follows from Lemma 1 and second from the fact that the link ij is sustainable at g. Since this is true for all such h_t, at each history σ does not agree with σ^*, it weakly lowers the expected payoff and σ^* has higher expected payoff compared to σ.

To prove the only if part, suppose that the inequality in Equation 3 doesn't hold for $ij \in g$ and g is a stable network. Then there exists σ such that σ, g is a bilateral equilibrium where all favors are provided and all links are kept. As in any bilateral equilibrium all favors are performed on the equilibrium path, the expected utility under any bilateral equilibrium is same as the expected utility under σ. Let $\tilde{\sigma}$ denote an alternative strategy such that $\tilde{\sigma}_i(h_{0,i}, j) = \{K, D\}$. By Corollary 8 in Appendix B, the expected utility difference between $\tilde{\sigma}$ and σ is

$$\frac{\alpha (1 - (1-p)^{d_j(g)})}{d_j(g)} \left[\frac{\delta}{1-\delta} u_i(g - ij) + c - \gamma - \frac{\delta}{1-\delta} u_i(g) \right] > 0 \tag{15}$$

where the inequality follows as ij is not sustainable. Thus $\tilde{\sigma}_i$ is a profitable deviation.

A.4. Proof of Theorem 1

Let i be a player who has n neighbors (who all have n neighbors) at g. For g to be stable, from Proposition 3,

$$-c + \gamma + \frac{\delta}{1-\delta} u_i(g) \geq \frac{\delta}{1-\delta} u_i(g - ij) \text{ for } j \in N_i(g) \tag{16}$$

Rearranging, we obtain

$$\begin{aligned}-c+\gamma+\frac{\delta}{1-\delta}\Big[\alpha v\Big(1-(1-p)^n\Big)-n\frac{\alpha}{n}c\Big(1-(1-p)^n\Big)\Big]&\\ \geq\frac{\delta}{1-\delta}\Big[\alpha v\Big(1-(1-p)^{n-1}\Big)-(n-1)\frac{\alpha}{n}c\Big(1-(1-p)^n\Big)\Big]&\end{aligned}\tag{17}$$

The equation highlights the trade-off between keeping the neighbor and incurring the cost and not incurring the cost but having less continuation payoff. Rearranging we obtain

$$-c+\gamma+\frac{\delta}{1-\delta}\left(\alpha v\left((1-p)^{n-1}-(1-p)^n\right)-\frac{\alpha c}{n}\left(1-(1-p)^n\right)\right)\geq 0\tag{18}$$

Equation 18 can be written this as a cutoff on c, denoting the highest possible value of c to cooperate when a player has n neighbors (who all have n neighbors):

$$c(n)=\frac{\delta\alpha v(1-p)^{n-1}p}{1-\delta+\frac{\alpha}{n}\delta[1-(1-p)^n]}+\gamma\tag{19}$$

First, notice that $\lim_{n\to\infty}c(n)=0$, so the number of people a player can cooperate is bounded for $c>0$. Now, we show $c(n)>c(n+1)$ for all n:

$$\begin{aligned}c(n)&=\frac{\delta\alpha v(1-p)^{n-1}p}{1-\delta+\frac{\alpha}{n}\delta[1-(1-p)^n]}+\gamma\\ &=\frac{\delta\alpha v(1-p)^{n-1}p}{1-\delta+\frac{\alpha}{n}\delta[p+p(1-p)+p(1-p)^2+\ldots+p(1-p)^{n-1}]}\frac{1-p}{1-p}+\gamma\\ &>\frac{\delta\alpha v(1-p)^{n}p}{1-\delta+\frac{\alpha}{n}\delta[p(1-p)+p(1-p)^2+\ldots+p(1-p)^{n}]}+\gamma\\ &>\frac{\delta\alpha v(1-p)^{n}p}{1-\delta+\frac{\alpha}{n+1}\delta[p+p(1-p)+p(1-p)^2+\ldots+p(1-p)^{n}]}+\gamma\\ &>\frac{\delta\alpha v(1-p)^{n}p}{1-\delta+\frac{\alpha}{n+1}\delta[1-(1-p)^{n+1}]}+\gamma=c(n+1)\end{aligned}\tag{20}$$

Where first inequality holds as we only partially multiply the denominator with $(1-p)<1$ and second inequality is true as $p>(1-p)^kp$ for all $k\geq 1$. As $c(n)$ monotonically decreases to 0, there is a B^* such that $c(n)>c$ for all $n\leq B^*=c(n)<c$ for all $n>B^*$.

To prove the first part, assume there is a stable network g where at least one player have more than B^* links. Let i be the player who has the most links at g. Then $d_i(g)=k>B^*$. Let $j\in N_i(g)$ and $N_j(g)=k'\leq k$. Then

$$\begin{aligned}\frac{\delta}{1-\delta}(u_i(g)-u_i(g-ij))-c+\gamma&=-c+\gamma+\frac{\delta}{1-\delta}\left(\alpha v\left((1-p)^{k-1}-(1-p)^k\right)-\frac{\alpha c}{k'}\left(1-(1-p)^{k'}\right)\right)\\ &\leq -c+\gamma+\frac{\delta}{1-\delta}\left(\alpha v\left((1-p)^{k-1}-(1-p)^k\right)-\frac{\alpha c}{k}\left(1-(1-p)^{k}\right)\right)<0\end{aligned}$$

where first inequality follows from $\frac{\alpha c}{k'}\left(1-(1-p)^{k'}\right)$ is decreasing in k' and second from the fact that $k>B^*$. Thus, by Proposition 3, g is not stable.

To prove the second part, assume at g all players have exactly B^* links. Then, $-c+u_i(g) \geq u_i(g-ij)-\gamma$ for all j and by Proposition 3 g is a stable network.

A.5. Proof of Theorem 2

Let g denote the network where all players have B^* links. Let $\{i,j\}=S$ and g' denote a network obtainable from g via deviations by S that violates strong stability.

Case 1: Either $d_i(g') > B^*$ or $d_j(g') > B^*$. Observe that the only possible link that exists in g' but does not exist in g is $\{ij\}$. Thus, if $d_i(g') > B^*$, then $d_i(g') = B^*+1$ and $d_j(g') \leq B^*+1$.

Lemma 2. *Suppose that $d_i(g) = B^*+1$ and $d_j(g) \leq B^*+1$. Then the link ij is not sustainable at g.*

Proof. For a contradiction, suppose that that ij is sustainable. Then

$$u_i(g)-u_i(g-ij) = \alpha v(1-p)^{d_i(g)-1}p - \alpha c\left(\frac{1-(1-p)^{d_j(g)}}{d_j(g)}\right) \geq c-\gamma \tag{21}$$

Moreover,

$$\alpha v(1-p)^{B^*}p-\alpha c\left(\frac{1-(1-p)^{B^*+1}}{B^*+1}\right) \geq \alpha v(1-p)^{d_i(g)-1}p-\alpha c\left(\frac{1-(1-p)^{d_j(g)}}{d_j(g)}\right) \geq c-\gamma \tag{22}$$

where the first inequality holds has $d_j(g) \leq B^*+1$ and the term $\frac{1-(1-p)^{d_j(g)}}{d_j(g)}$ is decreasing in $d_j(g)$. Let g' denote the network where all players have B^*+1 neighbors. Equation 22 implies that all links in g' are sustainable and by Proposition 3, g' is stable, which is a contradiction to the definition of B^*+1. □

By Lemma 2, the link ij is not sustainable at g', which is a contradiction.

Case 2: $d_i(g') \leq B^*$ and $d_j(g') \leq B^*$. As the only possible link that exists in g' but does not exist in g is $\{ij\}$, for all $k \in N_i(g') \cup N_j(g')$, $d_k(g') \leq B^*$.

Lemma 3. *Suppose that $d_j(g) = B^*$ for all $j \in N_i(g)$ and $d_i(g) < B^*$. If $d_k(g) = B^*$ and $ik \notin g$, then $u_i(g+ik) > u_i(g)$.*

Proof. Let $d = d_i(g)$.

$$\begin{aligned} u_i(g+ik)-u_i(g) =& \alpha v(1-p)^d p - \alpha c\frac{1-(1-p)^{B^*}}{B^*} \\ &> \alpha v(1-p)^{B^*}p - \alpha c\frac{1-(1-p)^{B^*}}{B^*} \geq \frac{1-\delta}{\delta}(c-\gamma) > 0 \end{aligned} \tag{23}$$

where the first inequality holds as $d < B^*$, and the second follows from the definition of B^*. □

Lemma 4. *Suppose both i and all $j \in N_i(g)$ has B^* links at g. Moreover, both i and all $j \in N_i(g')$ has (weakly) fewer than B^* links at g'. Then $u_i(g) \geq u_i(g')$.*

Proof. If i and all $j \in N_i(g')$ has B^* links, then $u_i(g) = u_i(g')$. If $d_i(g') = d_i(g)$ but there exists a $j \in N_i(g')$ with $d_j(g') < B^*$, then the result follows from Proposition 2. If $d_i(g') < d_i(g)$, define g'' as the network where $d_i(g'') = d_i(g')$ but $d_j(g'') = B^*$ for all $j \in N_i(g'')$. Then $u_i(g) > u_i(g'') \geq u_i(g')$ where the first inequality follows from Lemma 3 and the second follows from Proposition 2. □

Applying Lemma 4 for i and j, we have $u_i(g) \geq u_i(g')$ and $u_j(g) \geq u_j(g')$, which is a contradiction. This proves the first part of the theorem. To prove the second part, I first prove the following lemma.

Lemma 5. *In any strongly stable network and $k < B^*$, there can be at most $k + 1$ players with k links.*

Proof. Let g denote a strongly stable network. For a contradiction, assume for some $n < B^*$, there are $n' > n$ players with n number of links. Then at least two of those players are not linked, let i and j denote those players. Then $g' = g \cup \{ij\}$ satisfies all conditions for violation of strong stability and thus g is not strongly stable. □

As a result, at any strongly stable network, for any $n < B^*$ there can be at most n players with n links, which proves that total number of players with fewer than B^* can be at most $\frac{(B^*)^2 + B^*}{2}$, which is independent of N.

A.6. Proof of Proposition 4

Immediate from differentiating the LHS of equation 18 with respect to c, δ, v and γ.

A.7. Proofs of Proposition 5

Let $w(i, g) = \frac{1-(1-p)^{d_i(g)}}{d_i(g)}$ denote the probability i asks a favor from one of her partners. For a contradiction, suppose that (g, t) is strongly stable with transfers (SST) and there exists $i \in \tilde{N}$ with $d_i(g) < B^*$. Moreover, let i be one of the lowest degree players in $\tilde{N}$ at g, in other words, for all $j \in \tilde{N}$, $d_i(g) \leq d_j(g)$.

Fix j and k such that $j \in \tilde{N}$, $jk \in g$, $ij \notin g$, and $ik \notin g$.[33]. Let $g_j = g - kj + ij$ and $g_k = g - kj + ik$, and t_{ij} and t_{ik} the proposed transfers from i to j and k. Note that

$$u_j(g_j) - u_j(g) = w(j, g)t_{jk} + w(k, g_j)(c - t_{kj}) - w(i, g_j)(c - t_{ij}) \tag{24}$$

[33]Such j and k exist as $|\tilde{N}| \geq B^*$ and $d_i(g) < B^*$

$$u_k(g_k) - u_k(g) = w(k,g)t_{kj} + w(j,g_k)(c - t_{jk}) - w(i,g_k)(c - t_{ik}) \tag{25}$$

Set the transfers as follows, which makes sure j and k are indifferent between removing their relationship and adding a relationship with i.

$$t_{ij} = \frac{w(i,g_j)c - w(j,g)t_{jk} - w(k,g_j)(c - t_{kj})}{w(i,g_j)} \tag{26}$$

$$t_{ik} = \frac{w(i,g_k)c - w(k,g)t_{kj} - w(k,g_k)(c - t_{jk})}{w(i,g_k)} \tag{27}$$

Let t_j denote the transfer scheme where t_{ij} replaces the ij'th entry of t and t_k denote the transfer scheme where t_{ik} replaces the ik'th entry of t. I will now show that either i and j violate strong stability with transfers via g_j, t_j or i and k violate strong stability with transfers via g_k, t_k.

First, note they way t_{ij} and t_{ik} set makes sure that i and j are weakly better off and j's (k's) sustainability condition holds for the link ij (ik). Moreover, summing both sides and plugging in $w(j,g) = w(j,g_j)$ and $w(k,g) = w(k,g_k)$, we get

$$t_{ij} + t_{ik} = \frac{w(i,g_j) + w(i,g_k) - w(j,g_j) - w(k,g_k)}{w(i,g_j)}c \tag{28}$$

We now show that the i's sustainability condition holds for at least one of the links ij or ik at the proposed transfers. Sustainability conditions for i is given by

$$u_i(g_j) - u_i(g_j - ij) = \alpha(1-p)^{d_i(g)}pv - w(i,g_j)t_{ij} - w(j,g_j)c \geq \frac{1-\delta}{\delta}(c - \gamma) \tag{29}$$

$$u_i(g_k) - u_i(g_k - ik) = \alpha(1-p)^{d_i(g)}pv - w(i,g_k)t_{ik} - w(k,g_k)c \geq \frac{1-\delta}{\delta}(c - \gamma) \tag{30}$$

Summing both sides, plugging in Equation 28 and cancelling terms, we obtain that $u_i(g_j) - u_i(g_j - ij) + u_i(g_k) - u_i(g_k - ik)$ is equal to

$$2\alpha(1-p)^{d_i(g)}pv - 2w(i,\hat{g})c \tag{31}$$

where $w(i,\hat{g}) = w(i,g_j) = w(i,g_k)$. Thus, to show that at least one of the LHS of Equations 29 and 30 is greater than $\frac{1-\delta}{\delta}(c-\gamma)$, it is enough to show that

$$\alpha(1-p)^{d_i(g)}pv - w(i,\hat{g})c = \alpha(1-p)^{d_i(g)}pv - \alpha c\left(\frac{1-(1-p)^{d_i(g)+1}}{d_i(g)+1}\right) \geq \frac{1-\delta}{\delta}(c-\gamma) > 0 \tag{32}$$

which holds as $d_i(g) < B^*$. This shows that if $i \in \tilde{N}$, then $d_i(g) \geq B^*$. To finish the proof, suppose for a contradiction $d_i(g) = d > B^*$ for some i. Moreover, suppose that i is one of the players with the highest number of links at g.

Lemma 6. *If $ij \in g$, then $j \in \tilde{N}$ and $d_j(g) = B^*$.*

Proof. Let $d_j(g) = m$. For the link ij to be sustainable, it must be that

$$\alpha\left(v(1-(1-p)^{d-1})p + \frac{(1-(1-p)^m)}{m}(t_{ji}-c) - \frac{(1-(1-p)^d)}{d}t_{ij}\right) \geq \frac{1-\delta}{\delta}(c-\gamma) \quad (33)$$

$$\alpha\left(v(1-(1-p)^{m-1})p + \frac{(1-(1-p)^d)}{d}(t_{ij}-c) - \frac{(1-(1-p)^m)}{m}t_{ji}\right) \geq \frac{1-\delta}{\delta}(c-\gamma)$$

As $d > B^*$, $\alpha v(1-(1-p)^{d-1})p - \alpha c\frac{(1-(1-p)^d)}{d} < \frac{1-\delta}{\delta}(c-\gamma)$. Then Equation 33 implies that $t_{ji} > 0$ and $j \in \tilde{N}$. We have already shown that any member of $\tilde{N}$ has at least B^* links. Moreover, summing both sides and rearranging

$$\alpha v(1-(1-p)^{d-1})p - \alpha c\frac{(1-(1-p)^m)}{m} + \alpha v(1-(1-p)^{m-1})p - \alpha c\frac{(1-(1-p)^d)}{d} \geq 2\frac{1-\delta}{\delta}(c-\gamma)$$

As $d > B^*$, $\alpha v(1-(1-p)^{d-1})p - \alpha c\frac{(1-(1-p)^d)}{d} < \frac{1-\delta}{\delta}(c-\gamma)$, which implies that

$$\alpha v(1-(1-p)^{m-1})p - \alpha c\frac{(1-(1-p)^m)}{m} > \frac{1-\delta}{\delta}(c-\gamma) \quad (34)$$

However, this is a contradiction whenever $m > B^*$, thus $m = B^*$. □

Lemma 6 implies that there exists j and k such that $ij \in g$, $ik \in g$ and $jk \notin g$. To see why, if this does not hold, all neighbors of i must be linked, which implies that they have more than B^* links, contradicting Lemma 6. Moreover, as $ij \in g$, $ik \in g$ Lemma 6 implies that $d_j(g) = d_k(g) = B^*$. The following Lemma shows that j and k are strictly better off by removing their relationship with i and forming a link between themselves without transfers and finishes the proof. In Appendix C.2 I also prove that the set of SST networks is not empty.

Lemma 7. *Let $g' = g - ij - ik + kj$, $t'_{kj} = t_{jk} = 0$ and t' is otherwise equal to t. (g', t') violates strong stability with transfers at (g, t).*

Proof. First, note that, for the link ij to be sustainable, it must be that

$$\alpha\left(v(1-(1-p)^{d-1})p\frac{(1-(1-p)^{B^*})}{B^*}(t_{ji}-c) - \frac{(1-(1-p)^d)}{d}t_{ij}\right) \geq \frac{1-\delta}{\delta}(c-\gamma)$$

Rearranging, we obtain

$$\alpha\frac{(1-(1-p)^{B^*})}{B^*}t_{ji} - \alpha\frac{(1-(1-p)^d)}{d}t_{ij} \geq \alpha c\frac{(1-(1-p)^{B^*})}{B^*} + \frac{1-\delta}{\delta}(c-\gamma) - \alpha v(1-(1-p)^{d-1})p \quad (35)$$

We have the following

$$\hat{u}_j(g',t') - \hat{u}_j(g,t) = \frac{(1-(1-p)^{B^*})}{B^*}t_{ji} - \frac{(1-(1-p)^d)}{d}t_{ij} - \alpha c\frac{(1-(1-p)^{B^*})}{B^*} + \alpha c\frac{(1-(1-p)^d)}{d}$$
$$\geq \alpha c\frac{(1-(1-p)^d)}{d} + \frac{1-\delta}{\delta}(c-\gamma) - \alpha v(1-(1-p)^{d-1})p > 0$$

where the first inequality is obtained by plugging in Equation 35 and second from the definition of B^* and that $d > B^*$. Repeating the same steps, we can obtain $\hat{u}_k(g', t') - \hat{u}_k(g, t) > 0$. As $\hat{u}_j(g' - jk) = \hat{u}_j(g - ij)$ and $\hat{u}_k(g' - jk) = \hat{u}_k(g - ik)$ and both ij and ik are sustainable at (g, t), we have that the link jk is sustainable at g', t', which finishes the proof. □

A.8. Proof of Proposition 6

Let $\overline{B} = \max_{\nu \in \mathcal{V}} B(\nu)$. The following lemma derives an upper bound on the number of players who do not attain their cooperation bound.

Lemma 8. *For any $k < \overline{B}$ there can be at most $k+1$ players with k links who do not attain their cooperation bound.*

Proof. Let g denote a strongly stable network. For a contradiction, assume for some $k < \overline{B}$, there are $k' > k$ players with k number of links who do not attain their cooperation bound. Then at least two of those players are not linked, let i and j denote those players. Then $g' = g \cup \{ij\}$ satisfies the conditions of Definition 7 and g is not strongly stable. □

Lemma 9. *If two players attain their cooperation bound and are linked with a player who have fewer links than them, then they must be linked.*

Proof. Let i and j denote the two players who are linked with a player who have fewer links than them and who attain their cooperation bound. Let k_i and k_j denote these players ($k_i = k_j$ is possible). Then $g' = g \cup \{ij\} \setminus \{ik_i, jk_j\}$ satisfies the conditions of Definition 7 and g is not stable. □

If all players have types from $\mathcal{V}$, for any society with N members, there can be at most $\sum_{b=1}^{\overline{B}}(b+1)$ players who do not attain their cooperation bound, which does not depend N. Hence as $N \to \infty$, the fraction of players who attain their cooperation bounds (and fraction of players who have links with others who have fewer links than them) goes to 1.

A.9. Proof of Proposition 7

First, note that RHS of equation 19 is same for both rich and poor players. Thus, there exists c such that equation 19 holds with equality when $n = B^*(c_p) + 1$, which is $\tilde{c}_r$. To prove the first part, suppose that the inequality is small and g is SST. The result follows from the same steps in Lemma 21 in Appendix C.2, replacing c with c_r if $i \in N_r$ and with c_p if $i \in N_p$.

To prove the second part, suppose that the inequality is large and g is SST.

Lemma 10. *If $i \in N_r$, then $d_i(g) = B^*(c_r)$.*

Proof. Follows from replicating the steps in Proposition 5. ☐

Lemma 11. *If $i \in N_p$, then $d_i(g) < B^*(c_r)$.*

Proof. Suppose that there exists $i \in N_p$ with $d_i(g) \geq B^*(c_r)$. Without loss of generality, let i be the poor player with highest number of links. Take any $k \in N_i(g)$. We have the following

$$\alpha v(1-p)^{d_i(g)-1}p - \alpha c_p \frac{1-(1-p)^{d_k(g)}}{d_k(g)} \leq \alpha v(1-p)^{B^*(c_r)-1}p - \alpha c_p \frac{1-(1-p)^{d_k(g)}}{d_k(g)} \\ \leq \alpha v(1-p)^{B^*(c_r)-1}p - \alpha c_p \frac{1-(1-p)^{B^*(c_r)}}{B^*(c_r)} < \frac{1-\delta}{\delta}(c_p - \gamma) \tag{36}$$

where the first inequality holds as $d_i(g) \geq B^*(c_r)$, the second inequality holds as $d_k(g) \leq B^*(c_r)$ and third holds from the definition of $B^*(c_r)$ and the fact that $c_p > c_r$. Equation 36 implies the link ik is not sustainable, thus g is not stable, which is a contradiction. ☐

Lemma 12. *If $i \in N_r$ and $j \in N_p$, then $ij \notin g$.*

Proof. Suppose that $ij \in g$. Note that as $i \in N_r$ and $j \in N_p$, $t_{ji} = 0$.

Take a $k \in N_r$ with $k \notin N_i(g)$. Moreover, as $d_i(g) = d_k(g) = B^*(c_r)$, there exists m such that $km \in N_k(g)$ with $m \notin d_i(g)$. First, if $m \in N_p$, by Lemma 11, $d_m(g) < d_i(g) = B^*(c_r)$. Moreover, as $m \in N_p$, $t_{mk} = 0$. Thus, $g' = g - ij - km + ik$ with proposed transfers $t_{ik} = t_{ki} = 0$ violate SST. To show this, I first show that i is strictly better off and i's sustainability constraint is met.

$$\hat{u}_i(g', t') - \hat{u}_i(g, t) = \alpha \frac{1-(1-p)^{B^*(c_r)}}{B^*(c_r)} t_{ij} + \alpha c \left(\frac{1-(1-p)^{d_j(g)}}{d_j(g)} - \frac{1-(1-p)^{B^*(c_r)}}{B^*(c_r)} \right) > 0 \tag{37}$$

where the inequality holds as $d_j(g) < B^*(c_r)$. This also implies that the sustainability constraint of i for the link ik is satisfied at g', t' as $\hat{u}_i(g' - ik, t') = \hat{u}_i(g - ij, t)$ and the link ij was sustainable at g. Second, I show that k is weakly better off and k's sustainability constraint is met.

$$\hat{u}_k(g', t') - \hat{u}_k(g, t) = \alpha c \left(\frac{1-(1-p)^{d_m(g)}}{d_m(g)} - \frac{1-(1-p)^{B^*(c_r)}}{B^*(c_r)} \right) > 0 \tag{38}$$

where the inequality holds as $d_m(g) < B^*(c_r)$. This also implies that the sustainability constraint of k for the link ik is satisfied at g', t' as $\hat{u}_k(g' - ik, t') = \hat{u}_k(g - ij, t)$ the link km was sustainable at g.

Second, suppose that $m \in N_r$. Note that this also implies that $d_m(g) = B^*(c_r)$. Without loss of generality, let $t_{km} \geq t_{mk}$ (otherwise, replace m and k). Consider $g' = g - ij - km + ik$ with proposed transfers $t'_{ik} = t'_{ki} = 0$. We will show that this constitutes a violation of SST. The proof for showing that i is strictly better off and i's sustainability constraint is met is exactly same as the first case. For k, we have the following:

$$\hat{u}_k(g', t') - \hat{u}_k(g, t) = \alpha \left(\frac{1 - (1-p)^{B^*(c_r)}}{B^*(c_r)} \right) (t_{km} - t_{mk}) \geq 0 \tag{39}$$

where the inequality holds as $t_{km} \geq t_{mk}$. This also implies that the sustainability constraint of k for the link ik is satisfied at g', t' as $\hat{u}_k(g' - ik, t') = \hat{u}_k(g - ij, t)$ the link km was sustainable at g. □

The following lemma proves the second part of the proposition.

Lemma 13. *If $i \in N_r$, then $d_i(g) = B^*(c_r)$. If $i \in N_p$, then $d_i(g) \leq B^*(c_p)$*

Proof. By Lemma 12, at g, there is no link between players in N_r and N_p. Let g_p and g_r denote the collection of links between players in N_p and N_r. Then g is stable if and only if g_p is stable in N_p and g_r is stable in N_r. Thus, by Theorem 1, if $i \in N_r$, $d_i(g) \leq B^*(c_r)$ and $i \in N_p$, $d_i(g) \leq B^*(c_p)$. From Lemma 10, we have $d_i(g) = B^*(c_r)$, which proves the result. □

A.10. Proof of Proposition 8

Lemma 14. *Suppose that g is a community stable network and $ij \in g$. If $\phi(i) \neq \phi(j)$, then* $\max(d_i(g), d_j(g)) \leq B^*$.

Proof. For a contradiction, suppose that $\max(d_i(g), d_j(g)) > B^*$. Without loss of generality, let $d_i(g) \geq d_j(g)$. As in any community equilibrium all favors are performed on the equilibrium path, the expected utility under any bilateral equilibrium is same for all σ. Let $\tilde{\sigma}$ denote an alternative strategy such that $\tilde{\sigma}_i(\tilde{h}_{0,i}, j) = \{K, D\}$ and removes the link with j whenever a favor is not performed in the past, but otherwise same as σ.

From Lemmas 18 and 19 in Appendix B, the expected utility difference between two strategies is given by

$$\frac{\alpha(1 - (1-p)^{d_j(g)})}{d_j(g)} \left[\frac{\delta}{1-\delta} u_i(g) - c + \frac{\delta}{1-\delta} u_i(g - ij) \right] < 0 \tag{40}$$

where the inequality follows as $d_i(g) > B^*$ and $d_i(g) \geq d_j(g)$. Thus $\tilde{\sigma}_i$ is a profitable deviation and g is not a community stable network. □

To prove the first part, for a contradiction suppose $d_i(g) > B^*$ and $\phi(i)$ is a small community. As $\phi(i)$ is a small community, there exists a j such that $j \in d_i(g)$ and $j \notin \phi(i)$. By Lemma 14, this is a contradiction and the result follows.

To prove the second part, suppose that $d_i(g) > B^*$ and i has a link with j with $j \notin \phi(i)$. This contradicts Lemma 14 and thus g is not a community stable network.

A.11. *Proof of Proposition 9*

First, note that $U(n+1,\kappa) = \alpha v\Big(1-(1-p)^n\Big) - \alpha c\Big(1-(1-p)^n\Big) - n\kappa \equiv \beta(n) - n\kappa$. Observe that $\frac{\partial^2}{\partial n^2}\beta(n) = -(1-p)^n \log^2(1-p) < 0$ for all $p \in (0,1)$. Thus $U(n,\kappa)$ is strictly concave in n. Next, $\lim_{n\to\infty}\beta(n) = \alpha(v-c)$ and $\lim_{n\to\infty} n\kappa = -\infty$. Hence, $\lim_{n\to\infty} U(n,\kappa) = -\infty$. As $U(0,\kappa) = 0$, $U(n,\kappa)$ is either decreasing in $n \in [0,\infty)$ or $U(n,\kappa)$ has a unique interior optimum at some $n \in \mathbb{R}$. In the first case, the optimum is 0, while in the second case U is maximized at either the floor or ceiling of the interior optimum, which proves the first part. The fact that $B(\kappa)$ is decreasing in κ is immediate.

To compare community enforcement with pure bilateral enforcement, note that when κ is small, cooperation with any number of players can be supported at a very low cost, therefore community enforcement is optimal. Second, when κ is large, clearly the cost of maintaining the community links dominates the benefit of cooperation, therefore bilateral enforcement is optimal. The following lemma shows that if community enforcement is optimal for a κ, then community enforcement is optimal under any $\kappa' < \kappa$.

Lemma 15. *$U(B(\kappa),\kappa)$ is decreasing in κ.*

Proof. Note that $U(B(\kappa'),\kappa') \geq U(B(\kappa),\kappa') > U(B(\kappa),\kappa)$, where the first inequality holds from the optimality of $B(\kappa')$ under κ' and second holds as U is decreasing in κ. Thus community cooperation under κ' dominates bilateral cooperation. □

Therefore, there exists κ_P such that community enforcement results in a higher payoff than pure bilateral enforcement if and only if $\kappa \leq \kappa_P$. Finally, if community cooperation is optimal under κ, then $B(\kappa) > B^*$ since cooperation of B^* players can be sustained bilaterally.

To prove the second part, note that the utility under legal enforcement is given by

$$U(\gamma) = \alpha(v-c)\left(1-(1-p)^{B^*(\gamma)}\right) - C(\gamma) \tag{41}$$

First, note that the first term is bounded above by $\alpha(v-c)$. Second, as $\lim_{\gamma\to c} C(\gamma) = \infty$, there exists $\overline{\gamma}$ such that $U(\gamma) < 0$ for all $\gamma > \overline{\gamma}$. Let $\overline{N} = B^*(\overline{\gamma})$. For each integer $n < \overline{N}$, let γ_n denote the γ value that satisfies the inequality in Equation 18 with equality, which means that γ_n is the lowest γ such that players can sustain n links. Take an $n^* \in \arg\max_{n<\overline{N}} U(\gamma_n)$, which is non-empty. Let $\gamma^* = \gamma_{n^*}$, which is the optimal level of legal punishment and $U(\gamma^*)$ is the highest per period utility players can obtain under legal enforcement. If κ is small enough, then $n^*\kappa < C(\gamma^*)$, and community enforcement is more efficient than legal enforcement. Conversely, if κ is large enough (*e.g.*, $\kappa > \alpha v$), community enforcement is not optimal. Then Lemma 15 implies that there exists κ_L such that community enforcement results in a higher payoff than pure bilateral enforcement if and only if $\kappa \leq \kappa_L$.

For Online Publication

B. Preliminary Results

I first prove an important Lemma about the beliefs in stable bilateral equilibria. I will write $\sigma_i(h_{t,i}, j) = X$ if σ_i puts probability 1 to the action X.

Lemma 16. *Suppose that g, σ is a stable bilateral equilibrium, $ij \in g$ and at $h_{t,i}$, both i and j provided all favors to each other and the link ij is kept. Then $\sigma_i(h_{t,i}, j) = (K, C)$.*

Proof. For a contradiction, suppose that $\sigma_i(h_{t,i}, j) \neq (K, C)$ for some strictly positive probability. Let p_R denote the probability i plays R and p_D denote the probability i plays (K, D).

Let h_t denote any history that agrees with $h_{t,ij}$ and such that all favors are provided and all links are kept between all players (thus $h_{t,ij}$ is the bilateral history between i and j at h_t). Observe that under σ, h_t has strictly positive probability, say β. If $p_R > 0$, then with probability βp_R, a link is not kept at σ, which means that g, σ is not a stable bilateral equilibrium. If $p_R = 0$, then it must be that $p_D > 0$. Let $\beta_{ij} > 0$ denote the probability that j asks a favor from i at h_t. Then i refuses to provide the favor with probability $\beta\beta_{ij}p_D > 0$, which means that g, σ is not a stable bilateral equilibrium. This proves the result. □

Lemma 17. *If (g, σ, μ) is a stable bilateral equilibrium, then at any history, any player believes that all favors are provided and all links are kept apart from the ones that he has already observed otherwise.*

Proof. Take a player i and history $h_{t,i}$. Let H_t denote the set of all histories h'_t such that $h'_{t,i} = h_{t,i}$. Take any history $h'_t \in H_t$ such that there is a deviation between j and m with $i \notin \{j, m\}$. Note that for each t, there are finitely many such h'_t.

Let $t' \leq t$ denote the period where first deviation occurred between j and m in h'_t. Moreover, As (g, σ, μ) is a stable bilateral equilibrium, the strategies of all players only depend to their bilateral history. Therefore that any σ^k that converges to σ, the probability that there is a deviation between j and m at period t' converges to 0 as $k \to \infty$. Thus, $\lim_{k\to\infty} \mu^k(h'_t) = 0$. As there are finitely many such h'_t, the sum of probabilities attached to them also converges to zero. Therefore, at any history, any player believes that all favors are provided and all links are kept apart from the ones that he has already observed otherwise. □

Corollary 8. *Suppose that (g, σ, μ) is a stable bilateral equilibrium, $h_{t,i}$ is a private history where i has not observed a deviation and $ij \in g$. Consider σ' such that $\sigma'(h_{t,i}, j) = \{K, D\}$ but is otherwise same as σ. If in period t, j asks a favor from i and i is able to provide that* *favor, the difference between the continuation payoffs between σ and σ' is*

$$c - \gamma + \frac{\delta}{1-\delta}(u_i(g) - u_i(g - ij)) \tag{42}$$

Proof. If j asks a favor from i in period t, not providing the favor increases period t utility by $c - \gamma$, but causes link ij to break. By Lemma 17, i believes that the rest of the network is unchanged under μ, thus, by Lemma 16, future payoff is given by $\frac{\delta}{1-\delta}(u_i(g) - u_i(g - ij))$. □

These results also extend immediately to the setting with community enforcement.

Lemma 18. *Suppose that g, σ is a community equilibrium, $ij \in g$ and at $\tilde{h}_{t,i}$, both i and j provided all favors to each other and the link ij is kept at $\tilde{h}_{t,ij}$. Then $\sigma_i(\tilde{h}_{t,i}, j) = (K, C)$.*

Proof. Follows from precisely same steps as in the proof of Lemma 16. □

Lemma 19. *Suppose that g, σ is a community equilibrium, $ij \in g$ and $\phi(i) \neq \phi(j)$. Take a history $\tilde{h}_{t,i}$ such that i has observed a deviation only in his bilateral history with j. Then i believes that all favors are provided and all links are kept apart from what he has observed in his bilateral history with j.*

Proof. Follows from precisely same steps as in the proof of Lemma 17. □

C. Additional Results

C.1. Bounded Cooperation under General Favor Provision Matrices

Let F denote the finite set of different favor types and M denote the favor provision probability matrix where $p_{if} \in 0 \cup [\underline{p}, 1]$ for some $\underline{p} > 0$, in other words, the probability each player can provide each favor is either 0, or bounded below by some number.

The following proposition shows that the bounded nature of cooperation is due to the substitutability of the favors and not symmetry of the fully substitutable model.

Proposition 10. *For any $(N, \alpha, M, v, c, \delta, \underline{p})$, there is a $B(\alpha, \underline{p}, v, c, \delta)$ such that if there exists j with $d_j(g) > B(\alpha, \underline{p}, v, c, \delta)$ then g is not stable.*

Proof. Let $\hat{u}_i$ denote the expected per period utility of player i. Suppose that σ, g is a stable bilateral equilibrium under beliefs μ. For $ij \in g$, let $\omega(i, j, g)$ denote the probability j requests a favor from i and i is able to provide it at any given period under g.

Corollary 9. *Suppose that $ij \in g$. Consider σ' such that $\sigma'(h_{0,i}, j) = \{K, D\}$, and σ' is same as σ otherwise. The difference in continuation payoffs between σ' and σ*

$$\omega(i, j, g)\left[-c + \gamma + \frac{\delta}{1-\delta}(\hat{u}_i(g)) - \hat{u}_i(g - ij)\right] \tag{43}$$

Proof. First, unless j requests a favor from i and i is able to provide it, both strategies are equivalent and gives i the same expected payoff. Suppose that at period 0, j requests a favor from i and i is able to provide it. Under σ', i plays D. By Lemma 17, i believes that the rest of the network is unchanged under μ, and thus by Lemma 16, i gets a payoff of $-\gamma+\frac{\delta}{1-\delta}\hat{u}_i(g-ij)$. Under σ, j plays C, and by Lemma 17 i gets a payoff of $-c+\frac{\delta}{1-\delta}\hat{u}_i(g-ij)$. □

As σ, g is a stable bilateral equilibrium, it must be that

$$\frac{\delta}{1-\delta}(\hat{u}_i(g)) - \hat{u}_i(g-ij) \geq c-\gamma \tag{44}$$

Suppose that at g, i has n_f links who can provide favor f with positive probability. Let $\underline{n} = \min_f n_f$.

$$\begin{aligned}\frac{\delta}{1-\delta}(\hat{u}_i(g)) - \hat{u}_i(g-ij) &\leq \frac{\delta}{1-\delta}\left(\sum_{f\in F}\alpha v\frac{1}{|F|}(1-\underline{p})^{\underline{n}}p_{jf}\right)\\ &\leq \frac{\delta}{1-\delta}\left(\sum_{f\in F}\alpha v\frac{1}{|F|}(1-\underline{p})^{\underline{n}}\right)\end{aligned} \tag{45}$$

where first inequality holds the probability that j is pivotal is maximized when $n_f = \underline{n}$ and all neighbors of i can provide each favor with probability $\underline{p}$. Second inequality holds as setting $p_{jf} = 1$ maximizes the term in the brackets. Note that RHS of Equation 45 goes to 0 as $\underline{n}$ goes to ∞. Then there exists b^* such that

$$\left(\sum_{f\in F}\alpha v\frac{1}{|F|}(1-\underline{p})^{b^*}\right) \geq c-\gamma > \left(\sum_{f\in F}\alpha v\frac{1}{|F|}(1-\underline{p})^{b^*+1}\right) \tag{46}$$

This implies the following lemma.

Lemma 20. *If g is stable and $ij \in g$, there is at least one favor type f such that j can provide f with positive probability and the number of other players who can provide f is weakly fewer than b^*.*

Lemma 20 implies that i can have at most $|F|b^*$ links. □

Proposition 10 extends Theorem 1 to the case with a general favor provision matrix. Moreover, $B(\alpha, \underline{p}, v, c, \delta)$ does not depend on M or N and the cooperation in the society is bounded as long as the favors are substitutable.

C.2. The Set of Strongly Stable Networks with Transfers is Nonempty

Lemma 21. *Suppose that $d_i(g) = B^*$ for all i and $t_{ij} = 0$ for all i and j. Then g, t is SST.*

Proof. Suppose that $\{i, j\}$ violates SST of g via g' and t'. As $d_i(g') \leq B^*+1$ and $d_j(g') \leq B^*+1$, we will consider 3 cases.

Case 1: $d_i(g') \leq B^*$ and $d_j(g') \leq B^*$. As the only link in g' that is not in g is ij, both i and j and all their neighbors at g' has at most B^* links. If $t'_i = t'_j$, then $\hat{u}_i(g', t') \leq \hat{u}_i(g, t)$ and $\hat{u}_j(g', t') \leq \hat{u}_j(g, t)$ and both players are weakly worse off, which means that this is not a violation of SST. If $t_i > t_j$, then $\hat{u}_i(g', t') < \hat{u}_i(g, t)$, which means that this is not a violation of SST. If $t_i < t_j$, then replacing i and j in the last sentence shows that this is not a violation of SST.

Case 2: $d_i(g') = B^* + 1$ and $d_j(g') = B^* + 1$. Without loss of generality, let $t'_i \geq t'_j$.

$$\hat{u}_i(g', t') - \hat{u}_i(g' - ij, t') \leq u_i(g') - u_i(g' - ij) \tag{47}$$

From definition of B^*, $u_i(g') - u_i(g' - ij) < \frac{1-\delta}{\delta}(c - \gamma)$, which shows that link ij is not sustainable at g' and this is not a violation of SST.

Case 3: $d_i(g') = B^* + 1$ and $d_j(g') \leq B^*$. First, note that $t'_j > t'_i$, as otherwise the link ij would not be sustainable at g'. Let $\hat{t} = \alpha \frac{1-(1-p)^n}{n} t'_j - \alpha \frac{1-(1-p)^{B^*+1}}{B^*+1} t'_i$ and $n = d_j(g') \leq B^*$. For the link ij to be sustainable at g' for i, we need

$$\alpha v(1 - (1 - p)^{B^*})p - \alpha c \frac{1 - (1 - p)^n}{n} + \hat{t} \geq \frac{1 - \delta}{\delta}(c - \gamma) \tag{48}$$

For $\hat{u}_j(g', t') \geq \hat{u}_j(g, t)$, the following inequality must be satisfied

$$\begin{aligned} &\alpha v(1 - (1 - p)^n) - \alpha c \frac{1 - (1 - p)^{B^*}}{B^*}(n - 1) - \alpha c \frac{1 - (1 - p)^{B^*+1}}{B^* + 1} - \hat{t} \\ &- \alpha v(1 - (1 - p)^{B^*}) + \alpha c \frac{1 - (1 - p)^{B^*}}{B^*} B^* \geq 0 \end{aligned} \tag{49}$$

which is equivalent to

$$\begin{aligned} &\alpha v(1 - (1 - p)^n) - \alpha c \frac{1 - (1 - p)^{B^*}}{B^*} n - \alpha c \frac{1 - (1 - p)^{B^*+1}}{B^* + 1} - \hat{t} \\ &- \alpha v(1 - (1 - p)^{B^*}) + \alpha c \frac{1 - (1 - p)^{B^*}}{B^*}(B^* + 1) \geq 0 \end{aligned} \tag{50}$$

As for $n < B^*$, $\alpha v(1 - (1 - p)^n) - \alpha c \frac{1-(1-p)^{B^*}}{B^*} n$ is increasing in n, previous equation implies

$$\begin{aligned} &\alpha v(1 - (1 - p)^{B^*}) - \alpha c \frac{1 - (1 - p)^{B^*}}{B^*} B^* - \alpha c \frac{1 - (1 - p)^{B^*+1}}{B^* + 1} - \hat{t} \\ &- \alpha v(1 - (1 - p)^{B^*}) + \alpha c \frac{1 - (1 - p)^{B^*}}{B^*}(B^* + 1) \geq 0 \end{aligned} \tag{51}$$

Cancelling terms and re-arranging

$$\alpha c \frac{1 - (1 - p)^{B^*}}{B^*} - \alpha c \frac{1 - (1 - p)^{B^*+1}}{B^* + 1} \geq \hat{t} \tag{52}$$

Re-arranging Equation 48 and plugging in $n = B^*$ to minimize RHS,

$$\hat{t} \geq \frac{1 - \delta}{\delta}(c - \gamma) - \alpha v(1 - (1 - p)^{B^*})p + \alpha c \frac{1 - (1 - p)^{B^*}}{B^*} \tag{53}$$

For both of these equation to hold, we need

$$\alpha v(1-(1-p)^{B^*})p - \alpha c\frac{1-(1-p)^{B^*+1}}{B^*+1} \geq \frac{1-\delta}{\delta}(c-\gamma) \tag{54}$$

which contradicts the definition of B^*, thus no such $\tilde{t}$ exists. □

C.3. Heterogeneous Favor Values Depending on Provider's Identity

In this section, I will discuss how to extend the results in Section 4.2 to a setting where the value of favor depending on the identity of provider. When the value of the favor is determined by the provider's type, players i (as Player 1) and j (as Player 2) play the following favor exchange game when i asks a favor from j:

Table 5: Payoffs in the Favor Exchange Game with Heterogeneous Players

	C	D
A	$(v_j, -c_j)$	$(0, -\gamma)$

In the setting of Section 4.2, the identity of $j \in N_i(g)$ matters for i only through $d_j(g)$. This allowed us to compute the effective type of each type as a function of how many relationships each player can sustain and show that players do not have relationships with lower types. When the value of the favor depends on the identity of the receiver, the type of j affects i's value not only through $d_j(g)$, but also through $v_j(g)$.

I now consider the setting of Section 4.2 with heterogeneous values instead of costs, and where the value of the favor is determined by the provider: rich players can provide higher value favors, $v_r > v_p$. In that case, we obtain a sharper separation between the networks of rich and poor as even a small increase in v makes rich prefer only other rich players, even if both types can sustain same number of relationships. I assume that there are more than $4B^*(v_r)$ rich players.

Proposition 11. *If g is strongly stable with transfers, then all rich players have $B^*(v_r)$ relationships and are only linked with other rich players. All poor players have at most $B^*(v_p)$ relationships.*

Proof. Suppose that g is strongly stable with transfers. I first prove the following lemma.

Lemma 22. *There can be at most $B^*(v_r)$ rich players who does not reach the cooperation bound $B^*(v_r)$.*

Proof. Suppose that there are more than $B^*(v_r)$ such players. Then there exists i and j in N_r such that $ij \notin g$. Consider $g' = g + ij$.

$$\hat{u}_i(g+ij) - \hat{u}_i(g) \geq \alpha p(1-p)^{d_i(g)} v_r - \alpha w(i, g+ij)t'_{ij} - \alpha w(j, g+ij)c + \alpha w(j, g+ij)t'_{ji} \quad (55)$$

$$\hat{u}_j(g+ij) - \hat{u}_j(g) \geq \alpha p(1-p)^{d_j(g)} v_r - \alpha w(j, g+ij)t'_{ji} - \alpha w(i, g+ij)c + \alpha w(i, g+ij)t'_{ij} \quad (56)$$

where the inequality holds as this supposes all other neighbors provides v_r favors and are never paid. Summing both equations, we get

$$\begin{aligned}[\hat{u}_i(g+ij) - \hat{u}_i(g)] + [\hat{u}_j(g+ij) - \hat{u}_j(g)] \geq & \alpha p(1-p)^{d_i(g)} v_r - \alpha w(i, g+ij)c \\ & + \alpha p(1-p)^{d_j(g)} v_r - \alpha w(j, g+ij)c\end{aligned} \quad (57)$$

Note that as $d_i(g) < B^*(v_r)$, both lines of the RHS of Equation 57 is greater than $(c-\gamma)((1-\delta)/\delta)$. Which implies that either Equation 55 or Equation 56 is greater than $(c-\gamma)((1-\delta)/\delta)$. Wlog, suppose that it is Equation 55. Then setting $t'_{ji} = 0$, there exists t'_{ij} that makes both lines greater than $2(c-\gamma)((1-\delta)/\delta)$. This shows that ij violate SST. Thus, there are at most $B^*(v_r)$ rich players with fewer than $B^*(v_r)$ links. □

Let N' denote the set of rich players who have $B^*(v_r)$ links.

Lemma 23. *At most $B^*(v_r)$ players in N' has a link with a poor player.*

Proof. If there are more than $B^*(v_r)$ of such players, then there exists i and j in N' such that $ij \notin g$, $ik \in g$, $jm \in g$ and both k and m are poor players. As $d_i(g) = d_j(g) = B^*(v_r)$, $g' = g - ik - jm + ij$ with $t'_{ij} = t'_{ji} = 0$ violate SST. □

Let N'' of rich players who have $B^*(v_r)$ links and are only linked to other rich players. Note that as there are $4B^*(v_r)$, there are at least $2B^*(v_r)$ and thus there exists i and j in N'' such that $ij \in g$. Fix the identity of i and j for the rest of the proof. Wlog, suppose that $t_{ij} \geq t_{ji}$.

Lemma 24. *All rich players have $B^*(v_r)$ links.*

Proof. Suppose that m is a rich player with fewer links. Let $g' = g - ij + im$ with $t'_{im} = 0$. I will show that there exists t'_{mi} such that g', t' violates SST.

Note that $u_i(g' - im, t') = u_i(g - ij, t)$. Then we have

$$u_i(g', t') - u_i(g, t) = \alpha w(j, g)c - \alpha w(m, g')c + \alpha w(m, g')t'_{mi} \quad (58)$$

$$u_m(g', t') - u_m(g' - im, t') \geq \alpha(1-p)^{d_m(g)} p v_r - \alpha w(i, g)c - \alpha w(m, g')t'_{mi} \quad (59)$$

we need to show that there exists $\alpha w(m, g')t_{mi}$ such that first equation is positive (which also implies the sustainability condition of i) and second equation is greater than $(c-\gamma)(1-\delta/(\delta))$.

To see why this is true, sum both equations, plug in $w(j,g) = w(i,g)$ to obtain

$$\alpha(1-p)^{d_m(g)} p v_r - \alpha w(m,g') \geq (c-\gamma)(1-\delta/(\delta)) \tag{60}$$

where the inequality holds as $d_m(g) < B^*(v_r)$. This shows that all rich players have $B^*(v_r)$ links. ☐

Given Lemma 24, it is enough to show that no rich player has a link with a poor player. Suppose that m is a such player, who is linked to a poor player n. Consider $g' = g - ij + im$ with $t'_{im} = 0$ and $t'_{mi} = 0$. As $t_{ij} \geq t_{ji}$, the sustainability condition of i is satisfied and i is weakly better off. Moreover, as n is poor, $t_{nm} = 0$ while $t_{mn} \geq 0$. Then we have

$$u_m(g', t') - u_m(g' - im, t') > u_m(g, t) - u_m(g - mn, t') \geq (c-\gamma)(1-\delta/(\delta)) \tag{61}$$

where first inequality holds as the favors provided by i are more valuable than n and i requires favors less often, and second as g is SST. However, this shows that $g' = g - ij + im$ with $t'_{im} = 0$ and $t'_{mi} = 0$ violates SST and thus no rich player has a link with a poor player. ☐

C.4. Legal Enforcement and Population Size

In this appendix, $C(\gamma)$ represents the total cost of maintaining a legal system, and per player cost is given by $C(\gamma)/N$. Next proposition shows that legal enforcement is the optimal enforcement mechanism is large societies.

Proposition 12. *For each N, there exists $\gamma^*(N)$ that maximizes the per period expected payoff under legal enforcement. There exists $\overline{N}$ and $\overline{N}(\kappa)$ such that*

- *Legal enforcement gives players higher payoffs than pure bilateral enforcement if and only if $N \geq \overline{N}$.*
- *Legal enforcement gives players higher payoffs than community enforcement if and only if $N \geq \overline{N}(\kappa)$.*

Moreover, if $\gamma^*(N)$ and $\gamma^*(N')$ are maximizers and $N > N'$, then $\gamma^*(N) \geq \gamma^*(N')$.

Proof. The utility under legal enforcement is given by

$$U(N,\gamma) \equiv \alpha(v-c)\left(1-(1-p)^{B^*(\gamma)}\right) - \frac{C(\gamma)}{N} \tag{62}$$

Since $B^*(\gamma)$ is increasing in γ and is defined on integers and $C(\gamma)$ is continuous, $U(\cdot,\gamma)$ only has jump discontinuities, thus is discontinuous at most countably many points.

Lemma 25. *$U(\cdot,\gamma)$ is upper-semi continuous in γ.*

Proof. From Corollary 7, $B^*(\gamma)$ is increasing in γ. Since $U(\cdot,\gamma)$ is increasing in $B^*(\gamma)$, at each point $U(\cdot,\gamma)$ is discontinuous, it jumps up. Since $U(\cdot,\gamma)$ is continuous at any other point, is is usc. ☐

Since $\lim_{\gamma\to c} C(\gamma) = \infty$, and the first term of $U(N,\gamma)$ is bounded, there exists $\epsilon > 0$, such that $U(N,\gamma) < U(N, c-\epsilon)$ for all $\gamma > c-\epsilon$. Then the existence of an optimal $\gamma^*(N)$ follows from the compactness of $[0, c-\epsilon]$ and upper semi-continuity of $U(\cdot,\gamma)$.

To prove the second part, for a contradiction, assume that $N > N'$ but $\gamma^*(N) < \gamma^*(N')$, where $\gamma^*(N)$ and $\gamma^*(N')$ are maximizers of $U(N,\gamma)$ and $U(N',\gamma)$. From optimality of $\gamma^*(N')$, we have $U(N',\gamma^*(N')) - U(N',\gamma^*(N)) \geq 0$. Moreover

$$\begin{aligned}&(U(N,\gamma^*(N')) - U(N,\gamma^*(N))) - (U(N',\gamma^*(N')) - U(N',\gamma^*(N)))\\ &= -\left(\frac{C(\gamma^*(N'))}{N} - \frac{C(\gamma^*(N))}{N}\right) + \left(\frac{C(\gamma^*(N'))}{N'} - \frac{C(\gamma^*(N))}{N'}\right) > 0 \qquad (63)\\ &> 0\end{aligned}$$

where the inequality follows from $N' < N$ and $C(\gamma^*(N')) - C(\gamma^*(N)) > 0$ and implies $(U(N,\gamma^*(N')) - U(N,\gamma^*(N))) > 0$, contradicting the optimality of $\gamma^*(N)$.

To compare the two bilateral enforcement mechanisms, observe that $\max_\gamma U(\gamma, N) = U(\gamma^*(N), N)$ is increasing in N since $U(\gamma^*(N), N) \geq U(\gamma^*(N'), N) > U(\gamma^*(N'), N')$ when $N > N'$. The result then follows from observing the optimality of legal enforcement when $N \to \infty$ and bilateral enforcement when $N \leq B^*(0)$.

To compare the efficiency of legal enforcement and community enforcement, I first show that as $N \to \infty$, legal enforcement can approximate the first best payoff of $\alpha(v-c)$ arbitrarily.

Lemma 26. $\lim_{N\to\infty} \max_\gamma U(N,\gamma) = \alpha(v-c)$

Proof. Let ϵ be given. First, note that there exists $\hat{\gamma}$ such that $\alpha(v-c)(1-(1-p)^{\hat{\gamma}}) > \alpha(v-c) - \epsilon/2$. Next, observe that there exists $\hat{N}(\hat{\gamma})$ such that for all $N \geq \hat{N}$, $C(\hat{\gamma})/N < \epsilon/2$, which proves the result. □

Let $U(B(\kappa),\kappa)$ denote payoff at the optimal community size under κ. Clearly, $U(B(\kappa),\kappa) < \alpha(v-c)$. The result than follows from the previous lemma and the fact that $\max_\gamma U(\gamma, N) = U(\gamma^*(N), N)$ is increasing in N. □