A Theory of Community Formation and Social Hierarchy

Susan Athey
Stanford GSB

Emilio Calvano
Bologna U. and CSEF

Saumitra Jha
Stanford GSB*

December 2, 2021

Abstract

We analyze the classic problem of sustaining trust when cheating and leaving trading partners is easy, and outside enforcement is difficult. We construct equilibria where individuals are *loyal* to smaller groups– *communities*– that allow repeated interaction. *Hierarchies* provide incentives for loyalty and allow individuals to trust agents to extent that the agents are actually trustworthy. We contrast these with other plausible institutions for engendering loyalty that require inefficient withholding of trust to support group norms, and are not robust to coalitional deviations. In communities whose members randomly match, we show that social mobility within hierarchies falls as temptations to cheat rise. In communities where individuals can concentrate their trading with pre-selected members, hierarchies where senior members are favored for trade sustain trust even in the presence of proximate non-hierarchical communities. We link these results to the emergence of trust in new market environments and early human societies.

1 Introduction

A long-standing puzzle in economics concerns how individuals maintain trust despite short-term incentives to cheat. Almost every type of exchange involves some sort of moral hazard problem, whereby the individual providing a good or a service may shirk

* *athey@stanford.edu* (corresponding author); *emilio.calvano@unibo.it*; *saumitra@stanford.edu*. We would like to thank David Baron, Avinash Dixit, Avner Greif, Jim Fearon, David Laitin, Pedro Miranda, Ian Morris, Josiah Ober, Paul Seabright, Adam Szeidl and participants in the Berkeley development seminar, Harvard-MIT organizational economics seminar, Nobel Symposium on Foundations of Organization (Stockholm), 4eme JEC workshop (Paris), and across Stanford for helpful comments. Gleb Romanyuk provided much valued assistance. Athey thanks the Toulouse Network for Information Technology for support.

on quality or effort, or fail to deliver the good. An enormous literature seeks to use the theory of repeated games to suggest solutions to this problem. In long-term bilateral relationships, the breakdown of future cooperation provides an incentive for cooperation. However, in larger groups, the probability of future interaction with the same individual may be small, and even worse, it may be possible for a cheater to make choices (such as relocating) to actively avoid those he has cheated in the past.

In the face of such fundamental challenges to cooperation in large groups, several natural alternatives can be considered (in Section 2, we review the literature in more depth). The first, which do not consider in this paper, is the *rule of law*. Here, we focus on settings where legal institutions are limited in their ability to enforce trust. The second is *community enforcement*: all individuals in a society agree to jointly punish defectors. This type of approach has limited value when individuals have the opportunity to *cheat and leave*, and the community enforcement technology does not extend across all potential trading partners. A third class of alternatives recognizes the crucial role of *loyalty*: strategies that provide incentives for individuals to concentrate their trust relationships with known partners over time. Our paper fits into this category.

Several perspectives on loyalty exist. One is that individuals differ in their innate propensity for good behavior, and individuals attempt to signal their type by cooperating. Loyalty is engendered because new trading partners are not known to be honest, and the signaling process takes time. In this paper, we set aside unobserved individual heterogeneity, and instead analyze other forces that can support loyal relationships.

We build a very simple model of a large population playing bilateral trust games. We consider several assumptions that enable rich and realistic behavior to emerge. First, we allow players to form *communities*. Technically, we allow players to choose a location, a choice that limits their ability to play trust games to other players that also chose that location in that period. However, individuals can relocate in every period, capturing the idea that players can cheat and leave. Second, we consider various technologies for *record keeping* within a community. The record keeping is limited, however; all current members of a community can see how long others have attended that community, as well as a uni-dimensional *status* which depends on public randomization but not in any way on any individual's trading history.

Third, a key exogenous parameter of the model, which we refer to as *trading selectivity*, is the extent to which a subset of community members can serve all of the trading needs of the community. This parameter is motivated by differences in the kinds of trust exchanges that could occur in the real world. Taking the location analogy literally, in a town or a neighborhood, you sometimes encounter people randomly, and have the opportunity to create value through trust. A neighbor can hold the door for you, help you carry your bags, warn you of a pothole ahead, sign for your package and deliver it to you, or call a doctor if you faint on the street. These trust opportunities arise due to proximity, and individuals cannot easily make choices to only interact within a small set of established relationships. By choosing your location, you inherently choose to interact with community members. A second type of interaction entails more choice in the selection of partners. For example, an individual may need advice or information, or they may need to procure goods or small amounts of labor where quality is difficult to contract upon. Then, an individual might be able to choose to interact with only certain individuals in a community; in such cases some individuals might engage in more trade than others. In our model, the only publicly observed characteristics of individuals are the length of their attendance and their *status* (the outcome of public randomization that depends on their attendance) so it is natural to consider that high status individuals would be selected more often for trade. Our paper analyzes only two polar cases, full trading selectivity and no trading selectivity, where in both cases trade must take place within a community. We compare the types of equilibria that can be sustained and their robustness properties across these two polar cases, noting that most real-world environments entail a range of trust interactions including both types of trading selectivity as well as intermediate cases.

Using our model, we proceed in several steps. First, we examine equilibria of a type that have been considered in the literature in the past in the context of our model. In these *identity investment equilibria*, individuals are forced to make costly (and wasteful) investments of either time or resources in order to start trading in a community. These investments may take a number of forms, including *culture-specific* investments or the costs of traveling to a specific physical location. Identity investment equilibria serve as a benchmark for the novel equilibria we introduce in our paper. Though natural, we critique these equilibria on several grounds. First, focusing on the *no trading selectivity* case, individuals artificially withhold trust from certain agents. They fail to trust certain individuals and thus lose out on valuable opportunities to create surplus, not because the individuals cannot be trusted (that is, not because it is impossible for their *cooperate* rather than *cheat and leave* incentive constraints to be satisfied), but rather to uphold a social convention. Furthermore, upholding that convention does not benefit a community itself: the community as a whole would create more expected surplus if it abandoned the convention of requiring new members to make investments. An individual community sustains trust because *other* communities impose entry costs, not because they play any role in the community itself. Thus, a coalitional deviation where all members of a community abandoned the initial investment requirement would benefit the community. On the other hand, if any single community did that, trade would break down in other communities.

This motivates us to propose several alternative structures for equilibria that sustain cooperation. We begin with the no trading selectivity case. There we develop what we call a *maximal trust hierarchical equilibrium*. In this equilibrium, when a trading opportunity arises between two individuals, trust–which in our model can occur at different scales– is always *maximal*: individuals trust one another up to the point where an agent is just indifferent between working and a cheat and leave strategy. The equilibrium is also *hierarchical*: agents with higher *status* are trusted more. As we will develop in the paper, a key feature of the equilibrium is that advancement to higher status– which can be thought of as *social mobility*– is probabilistic.

Once we have established conditions under which maximal trust hierarchical equilibria exist, we analyze more deeply the structure of the equilibria, as well as compare different equilibria in the class. We analyze the distribution of income, and look at tradeoffs between equality and efficiency.

Next, we turn to analyze the *trading selectivity* case. In this setting, we look at equilibria where trades are concentrated on agents with higher status.

A key tradeoff we emphasize is trading selectivity allows senior members of a hierarchy— who are trusted more— to potentially benefit from more trades. However, trading selectivity also implies that those who cheat others can simply avoid those that they have cheated. It is not ex ante obvious therefore whether cooperative equilibria with trading selectivity exist, let alone are robust.

It is also worth mentioning that analyzing these equilibria is not easy, since the incentives for arbitrary population sizes N depend on the realization of an N-dimensional state variable. However, we develop an intuitive way to analyze the trading selectivity case by introducing a parameter measuring the maximum number of tasks that individuals can perform (i.e., capacity constraints). We develop analytical solutions for a limit case, and then use numerical analysis to show that our results are not specific to that case.

We show not only these equilibria exist but that communities that employ these structures are robust in an interesting and novel way. Even if another community adopts a more *egalitarian* structure, individuals will still choose to join the communities with a hierarchical structure. This contrasts with identity investment equilibria, where the existence of a community that trusts newcomers undermines trust in other communities.

After developing the theoretical results, we then relate these results to three real-world applications. We apply our results to understanding the puzzling emergence of hierarchies among early human societies, to the development of impersonal citizenship, illustrated by the Roman Republic, and to the sustenance of trust in settings as different as marketplaces in modern Nigeria and in online communities. Though separated by time and geography, all our examples share the commonality of a challenging contracting environment, with increasing population sizes and the possibility of cheating and leaving for alternative venues. We trace how hierarchies emerged in each environment, and how hierarchies may be interpreted as providing a means of facilitating trust.

2 Existing Theoretical Perspectives

Our paper builds upon literatures on the importance of trust in economic development, on social networks and in historical political economy. The issue of cooperation based on reciprocity has attracted the attention of both social scientists (e.g. Kranton, 1996b) and evolutionary biologists (see Nowak and Sigmund (2005) for a survey.) The problem of sustaining trust in particular has long been seen as a fundamental question in economic development, and economics more generally (Arrow, 1974, McMillan, 2002).

Beyond the classic folk theorems, theorists traditionally focus on two types of mechanisms that overcome the trust problem: those that signal reputations and those that require third party enforcement. In reputation-based models, players learn about the others' type as honest or opportunistic as the relationship progresses. Equilibria in which more senior individuals are accorded more trust emerge as a result of these inferences (eg Sobel, 1985, Watson, 1999)).

Our formulation exhibits an analogous equilibrium dynamic without any unobserved individual heterogeneity. We propose an institution for supporting cooperation which, to the best of our knowledge, has not been proposed before. It has two key robustness advantages. First, hierarchical structures allow us to construct equilibria with binding incentive constraints for all individuals. As in Ghosh and Ray (1996)'s seminal 'starting small' paper, this means that in equilibrium no pair of players can jointly gain by deviating from the institutions. Beyond this, however, we show that hierarchical communities would survive even if other communities were to adopt a more egalitarian structure, while the reverse is not true— the presence of hierarchical communities can undermine cooperation in egalitarian communities. This helps explain why hierarchies are common. We also differ from an important line of research on cooperation in groups that employs multi-lateral enforcement among delimited coalitions, following Greif (1993) and Kandori (1992), by examining settings where multi-lateral enforcement is not possible and populations can be arbitrarily large.[1]

Dixit's (2003) study of trade expansion and enforcement is similar in motivation to ours. He uses a circular world as a geographical analogy for the costliness of information flows across distances, and examines how much cooperation can be sustained as the circle grows in size. He finds that small and large worlds can sustain greater trust than their intermediate counterparts. In small worlds, partners to a transaction are likely to know third parties in common, and thus are able to share information about defectors. In large worlds, developing a legal system becomes economical. In contrast, our model focuses on the case where information sharing about behavior within bilateral relationships is not possible.

Beginning with Klein and Leffler (1981), the role of specific investments or "cultural capital" has assumed an important role in research into trust. For example, Iannaccone (1992) applies the notion of specific investments to cults that provide club goods. There is an incentive to free-ride upon others' zealousness. In order to limit participation to the truly committed, religious practices, such as stigma and self-sacrifice develop to act as screening devices. Similarly, Fryer (2002) allows for identity-specific investments among African Americans and compares their effects on within-market trust with that of investing in general human capital or "acting white."[2] These works, however, take both

[1] For example, Woodruff (1998) finds that Mexican footwear manufacturers were also able to maintain trust through third party enforcement mechanisms and information sharing. In a manner analogous to Greif (1993)'s classic example of the Maghribi medieval traders, such trust is supported by multilateral enforcement among small, culturally homogeneous groups, underpinned by the threat of ostracism from the community or business coalition.

A feature of such identity-based mechanisms, as we will show, however, is that such groups fail to support cooperation when group sizes are large, there is limited information sharing that prevents third party enforcement and where the availability of alternative trading partners make it easy to cheat and leave, and thenceforth avoid the cheated party.

Furthermore, even when there some random chance of re-encountering a cheated party, such mechanisms require high degrees of coordination to sustain– coordinated barriers to entry in each group are raised to prevent cooperation failing in *other* groups. The relaxation of such barriers by any group coalition will lead to a failure in cooperation in all. Indeed, as Woodruff describes, third party enforcement and cooperation among Mexican footwear manufacturers appeared to break down as new opportunities to trade with the US emerged with trade liberalization, allowing alternative trading partners outside the coalition.

[2] In using the term "identity" we follow Akerlof and Kranton (2000). We differ, however, in the form that identity takes. In their formulation, group "identity" enters into individuals' utility functions. These identities and the associated "prescriptions" for behavior result in individual and group sanctions for violators of the group "code of conduct."

the set of identities and prescriptions to be exogenous.

The intuition underlying the identity-investment equilibrium– using barriers to entry such as costly "gifts" into new relationships to sustain cooperation in old relationships– has been noted by a number of important studies (eg Kranton, 1996a, Carmichael and Macleod, 1997, Ramey and Watson, 2001)). The focus of these studies has however chiefly been two-agent partnerships rather than broad groups.

Research has also begun to examine the role of such conventions for sustaining cooperation in groups. The role of membership fees in engendering loyalty to "insiders" with whom such costs have not been incurred has been explored by Board (2008), and the role played by time in acting as such a membership fee has been explored by Friedman and Resnick (2001) in the context of internet chat rooms. Sobel (2006) analyzes a model where individuals in a large population form bilateral relationships. He contrasts relational contracts with formal contracting as mechanisms for sustaining trust. As in our model, it can be inefficient for partnerships to be exclusive in every period. Sobel's model has relationships that permanently "grow stale."In his model, relationships based on relational contracts may last inefficiently long, because the institutions that support cooperation must entail costs of starting new relationships. Similarly, in the literature on cooperation in social networks (eg Jackson, 2003, Bloch, Genicot, and Ray, 2008), the concept of hierarchy is tied to network centrality that is based upon upon past trades or ties that link specific individuals.

In contrast in our hierarchical equilibrium, the social hierarchies we construct are "impersonal" in the sense that the actual agents in the hierarchy can change but they inherit the incentives of their rank and thus fully efficient exchange can be sustained. Instead, akin to military custom or the ideal type of a Weberian bureaucracy, individuals in our hierarchical equilibria *salute the rank* of others even if they have never encountered them before. This is a distinct advantage of social hierarchies that has not to our knowledge been explored in the economics literature.[3]

[3]The key distinction between an agent's rank and their personal identity naturally has a long tradition in sociology. In particular, we build on and further work, at least as early as White (1970), on mobility in hierarchical organizations and "chains of vacancies" created by openings at higher ranks of hierarchies (see also Gibbons (2005)). Specific identity investment also has parallels in an important literature, beginning with Kreps (1990), on the role of "corporate culture". Culture can create value for the firm through variety of means, including reducing costs of coordination and communication and improving commitment by managers (see also Hermalin (2010).) Homogeneity within firms can happen through selection and through learning and indoctrination (Van den Steen, 2005) which can also be thought of as specific investments in the firm culture or identity. Thus, though we abstract from the potential productive roles of identity investment, our discussion of the relative robustness of hierarchies and identity investment equilibria can shed light on the relative robustness of trust-enhancing aspects of

Thus, cooperative equilibria based upon the existence of barriers to form new relationships occupy a prominent role in theories of trust.[4] An important focus of our study is to analyze the robustness properties of such identity investment equilibria in environments where some groups adopt social hierarchies and there may be insurgent egalitarian groups that lower barriers to entry entirely.[5] Our analysis also differs from much of the existing literature on trust in its focus on endogenous group formation, hierarchical structures, and the problems associated with increasing population size.[6] It links and contributes instead to important literatures looking at cultural transmission (eg Boyd and Richerson, 1994, Bisin and Verdier, 2011, Doepke and Zilibotti, 2013) and the origins of hierarchy and formation of state-like institutions (Bates, Greif, and Singh, 2002, Besley and Persson, 2009, North, Wallis, and Weingast, 2009, Bowles and Choi, 2013, Seabright, 2013, Dow and Reed, 2013, Boix and Rosenbluth, 2014, Mayshar, Moav, Neeman, and Pascali, 2015), as we describe below.

3 The Model

We take as our departure point the classic Shapiro and Stiglitz (1984) model of moral hazard and unemployment in the job market, a useful benchmark for examining trust in trade relationships (Greif, 1993). Specifically, we consider a dynamic equilibrium model with individuals repeatedly playing a bilateral trust game with randomly matched partners that can be looked upon as an extended version of the classic prisoner's dilemma. These repeated interactions take place within *communities* that can be thought of as geographical locations (for example villages), virtual online groups or simply physical groups of people who recognize each others' affiliation. At every point in time individuals are exclusively assigned to one community. Crucially, in each period, they may leave a community for good and join a new one. With this sketch in mind, we now proceed to

corporate culture as well.

[4]A notable exception to the focus on entry barriers to sustain trust is work by Lindsey, Polak, and Zeckhauser (2003), who incorporate the notion of itinerant temptations into their study of long-term bilateral, exclusive relationships between individuals where there is "free love": no barriers to a new start. Existing relationships gain value the longer they exist. This makes them robust to break-ups. However, this occurs for different reasons in our model: in the social hierarchies we construct, senior agents engender more trust.

[5]In comparing the competitiveness of alternative forms of social organization, our paper has natural links to research in organizational ecology (e.g. Hannan, Polos, and Carroll (2007)) as well as reciprocity in traditional societies exposed to the market (Kranton, 1996b).

[6]The role played by subgroup defection in limiting patterns of cooperation in larger communities has also been studied by Genicot and Ray (2003) in the context of risk-sharing. We abstract from the risk-sharing advantages of having larger groups in our study.

introduce the formal equilibrium model.

There is a number of discrete time periods $t = 1, 2, ..., \infty$, a stationary population of P players and a finite set of communities of cardinality M of equal size $N = P/M$, with N integer and even. At the end of each period individuals survive with independent probability δ. At the outset of the game, individuals are randomly assigned to one community. The within period timing is as follows. Inside each community, individuals are randomly matched in principal-agent pairs. Each individual is principal to one agent and agent to one principal.[7] First, principals choose how much to trust the agent they are matched with, offering a scale of trade $\lambda \in [0, 1]$. Then, agents choose whether to work or shirk as a function of trust offered to them λ. Finally, observed the outcome of their interaction, all individuals choose whether or not to leave the community for good. If an agent works, payoffs are $\alpha\lambda w$ and $(1 - \alpha)\lambda w$ to the principal and agent, respectively. If an agent shirks, the agent receives a private benefit λs while the principal incurs a private cost $c \geq 0$. We assume $s > w > 0$ and $w > s - c$, so shirking is inefficient. To improve the exposition, we assume that α is positive albeit arbitrarily small. Formally this is $\alpha = 0$. This gives principals an incentive to trust *as much as possible* if they conjecture the agent works. If individuals leave, conditional on surviving, they are randomly assigned to another community inheriting the index of an individual who either died or left.[8] Remaining vacancies are filled with newborn individuals to keep the population stationary.

Players observe personal bilateral histories between themselves and any other player. These bilateral histories are reset whenever either individual dies or moves. So, new members start with a clean slate. Each community is equipped with a public randomization device and we assume that at the beginning of each period, community members observe its realization. We will provide more details when those are used, and in environments where public randomization is payoff-irrelevant, we simply omit reference to the realization of these devices, implicitly assuming that players simply ignore them. For the moment, simply note that they are meant to capture the idea that within each community there is a hierarchical structure, with individuals being assigned some 'status' which is publicly observable and that this 'status' evolves over time. Note that this

[7] So, in each community, N different pairs are formed every period, with each individual playing *both* roles. This assumption is meant to simplify the exposition, as we won't have to distinguish between the payoffs and incentives of different roles at the interim stage.

[8] We assume that P is so large that the probability that at least one individual did not survive in one of the other communities is so large that moving to a different community is always an option. Formally, when we consider incentive constraints, we ignore the possibility that other communities cannot accommodate an agent who seeks to leave his current community. As we shall see, this is a *conservative* treatment of the incentive constraints.

randomization device does not add anything to the observability of trading histories, but instead simply introduces a public history with public randomization.

A *period trading strategy* is a mapping from the set of all possible bilateral trading histories, and set of all possible realizations histories of the public randomization devices to the following choices: (i) as principal how much to trust each agent, (ii) as agent whether to shirk or work for each potential principal as a function of the trust level granted and whether to stay or leave.

As will be clear when we introduce our candidate strategy profiles, in what follows we focus on incentive provision through reciprocity. Specifically we focus on strategies that prescribe shirking if and only if the same two individuals meet again in the period after a deviation by either took place. By limiting our attention to punishments that last one period, rather than grim trigger or other more extensive punishments, we naturally make it *harder* to support cooperation for a given level of patience and group size.[9] We will use the same simplification in all of the different strategy profiles we study, allowing us to compare outcomes across different types of equilibria faced with similarly challenging conditions for enforcing contracts.

'Institution-Free' Equilibria

To benchmark the importance of institutions, we first consider classes of candidate sequential equilibria that are *institution-free* in the sense that individuals condition only on their bilateral trading history when determining trust. Not surprisingly, the combination of repeated interaction and a sufficiently high degree of patience enables cooperation despite a large number of individuals and alternative trading partners.

Consider institution-free strategies, which specify that all traders move every period. That is as if players were uniformly randomizing over communities in every period. So, they perceive it as equally likely to be matched with all other individuals in the population, as if there were no communities at all. On the equilibrium path, principals maximally trust, and agents work if and only if they are maximally trusted. Deviations from the strategy above at any time t trigger a one period reversion to the static equilib-

[9] This simplification also allows us to avoid a technical complication: when players punish one another by refraining from trade forever, in some cases we analyze they might contemplate fairly complex strategies, where they accumulate a set of "enemies" in one community (principals they cheated in the past) before leaving for another community. By focusing on establishing the parameter space for which equilibria exist even with weaker (one-period) punishments, we can ensure that they will also exist with stronger sanctioning of deviations as well.

rium of shirk and zero trust in period $t+1$ if both trading partners are matched again to each other. It is straightforward to show that these strategy profiles constitute an equilibrium if and only if:

$$s - w \leq \frac{\delta^2}{P-1}(w+c). \tag{1}$$

The left hand side is the maximum short term gain from deviating. The right hand side is the expected lost stream of surplus due to the breakdown of bilateral cooperation in $t+1$. It amounts to the lost payoff from serving as an agent w and the cost for being cheated as a principal c, discounted by the probability that both individuals survive in the next period, multiplied by the (subjective) probability that the two individuals will be randomly matched to one another in the sub-game starting after the deviation occurred.[10]

Equilibria Based On Identity Investments

As discussed above, one way of sustaining cooperation as populations grow is to require individuals to make *specific* investments in a cultural identity that marks them as members of a community. Such investments can take a number of forms, depending on the exogenous dimensions available through which groups may assert distinction between one another and conformity within themselves (e. g. ethnic or racial markers, language, culture, physical location, religious or ethical codes). Such investment can be either publicly observable or simply required to join existing communities. Suppose that in order to start afresh in a new group individuals are required to make a (wasteful) investment m[11]. as follows. On the equilibrium path, agents are fully trusted ($\lambda = 1$), work when trusted and return to the same community. Deviations from the strategy above at any time t trigger a one period reversion to the static equilibrium of shirk and zero trust if and only if both trading partners return and are matched again in $t+1$. Cheated principals always return (so that they are there to provide the punishment).

[10]Note that the condition above could have been maximally weakened had we considered equilibrium profiles in which a one time deviation leads to a 'grim trigger'– an infinite replication of the stage-game equilibrium with shirking and no trust forever thereafter. In that case, the right side of (1) should be replaced by $\frac{\delta^2}{1-\delta^2}\frac{1}{P-1}(w+c)$, i.e. the lost stream of future trading surplus due to the permanent loss of reciprocal trust.

[11]m is assumed weakly lower than the maximum payoff attainable $\frac{w}{1-\delta}$ as otherwise it wouldn't be individually rational to join any group.

Proposition 1. *An equilibrium in identity investment strategies exists if and only if*

$$s - w \leq \delta m, \tag{IC_L}$$

$$s - w \leq \frac{\delta^2}{N-1}(w + c). \tag{IC_R}$$

Constraint (IC_R) provides incentives to cooperate *within* groups. It is the analogue of (1) at the group (rather than at the population) level. It says that groups should be small enough. (IC_L) requires the short term gain to be lower than the future loss incurred due to individuals migrating to different groups.

The minimum investment required to guarantee cooperation is $\underline{m} = (s - w)/\delta$, and cooperation can be sustained for lower degrees of patience as the groups shrink in size, with minimal patience required for groups of size two: there are no gains from having larger groups. If $s > w/(1 - \delta)$ then the minimum investment $\underline{m}$ would exceed the maximum present gain from cooperation $w/(1 - \delta)$. Thus a necessary condition for cooperation to be supported is that $s \leq \frac{w}{1-\delta}$.

In an identity investment equilibrium, a reduction of distance between communities may make the assertions of ethnic or cultural differences become *more* desirable as they allow the creation of entry barriers and investments that sustain cooperation. Thus the need to sustain cooperation within communities can lead to increased cultural separation between them. Indeed, Fryer (2002) provides a compelling account of the costs of African Americans of being perceived to be "Acting White" in mixed neighbourhoods, and how overt cultural markers of identity are used to sustain cooperation in those communities.[12]

Though identity investment equilibria have real-world analogues, it is important to note that they also have several unattractive robustness features. First, they are not robust to coalitional deviations of a particular kind. Specifically, the role of the investments required by one group is to sustain cooperation in *other* groups. The fact that other groups require investments to join makes it unattractive to cheat and leave. But a given community could simply stop requiring investments, and it would not be affected as long as all other communities do. And indeed, if an individual arrived in a community and did not make the required investment, the identity investment strategies require individuals to withhold trust from the individual, incurring a loss of potential surplus, even

[12] It is also possible to study a closely related set of strategies, whereby players refrain from trading with newcomers when matched with them, or reduce their level of trade with newcomers (as in Friedman and Resnick (2001)). The incentives created by this type of *initiation* requirement are similar, but they also impose costs on senior members of a community who are matched with newcomers, creating additional inefficiency.

though that individual could be deterred from cheating given that all other communities require investments.

This highlights the second unattractive feature of this class of equilibria: the strategies require withholding trust from *trustworthy* individuals. By this we mean that the individual would find it more profitable to work when trusted than to cheat. Indeed, the individual can only be induced to cheat when trusted in this scenario (as required by the identity investment strategies) by the threat of withholding future trust if the player does not cheat. In the rest of the paper, we develop alternative equilibria that also induce loyalty, but that address at least the latter concern, and in some cases the former concern as well.

Hierarchical Institutions

In what follows we define a hierarchical structure as a set of labels (referred to in what follows as 'seniority levels') and a stochastic *advancement process* that describes how these labels evolve over time. Formally, let the interval $[0,1] \subset \mathbb{R}_+$ be the set of seniority levels and $G : [0,1] \times [0,1] \to [0,1]$ be a family of cumulative distribution functions. $G(\tilde{l}, l)$ is the subjective probability that an individual with seniority level l (inherited from the previous period) attaches to transitioning to seniority lower or equal than $\tilde{l}$, conditional on surviving. Assume that:

$$G(\tilde{l}, l') \leq G(\tilde{l}, l) \qquad \text{for all } \tilde{l},\ l' > l. \tag{A1}$$

$$G(\tilde{l}, l) > 0 \quad \text{if and only if } l \geq \tilde{l}, \tag{A2}$$

$$G \text{ continuous and twice differentiable in both arguments.} \tag{A3}$$

(A1) implicitly defines a partial order over seniority levels. That is, $G(\tilde{l}, l)$ describes a family of stochastic processes parametrized by l with an intuitive property: for any l, individuals with a relatively higher l have at least as high a probability of transitioning to a label greater than l. In other words, (A1) says that individuals with higher seniority level enjoy an advantage over lower ranked individuals in their prospects for promotion to a rank further up the ladder. (A1) is definitional, capturing the essence of hirarchical structures. (A2) restricts the advancement process between levels to only allow individuals to move upward. ($A3$) disallows discontinuous or non-smooth jumps in the probability of promotion. (A2) and ($A3$) are imposed to simplify the analysis. A discussion on robustness is provided at the end of the section. Let $g(\tilde{l}, l)$ denote the corresponding probability density function.

We specify one relatively weak institutional requirement: that communities keep a *public record* of members' seniority levels.[13] At the beginning of every period new seniorities are drawn according to the above process and the public record is updated. For those individuals with no previous record (newcomers), their initial level is drawn according to $G(l, 0)$. That is, it is as if their previous period seniority were zero.

Define *maximal trust hierarchical strategies* as follows: on the equilibrium path, principals trust l types at scale l, agents work and always return. All $l < 1$ agents are indifferent between working and shirking, so trust is maximal.[14] The punishment phase triggered by a deviation by either individual in t and conditional on the agent returning is a one period Nash reversion to no trust/shirk within the bilateral trading relationship. In the continuation game after any one stage deviation to the above, the principal always returns and the agent always leaves.

A *maximal trust hierarchical equilibrium* (abbreviated HE) is a function G and a set of *maximal trust hierarchical strategies* such that given G, the strategies constitute a sequential equilibrium of the underlying game.

If a HE exists, an individual of seniority l's period payoff along the equilibrium path is simply lw. Let $v(l)$ denote the expected discounted equilibrium payoff for an individual of previous period seniority l right before the random advancement process takes place. $v(l)$ by definition should satisfy:

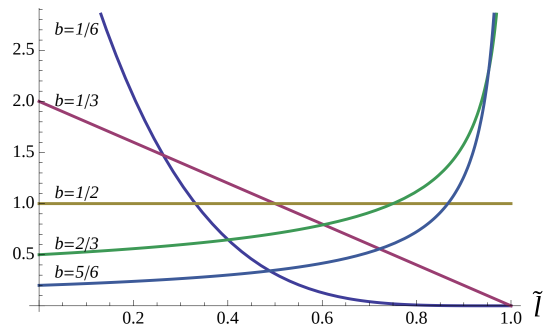
$$v(l) = \begin{cases} \int_l^1 g(\tilde{l}, l)(\tilde{l}w + \delta v(\tilde{l}))d\tilde{l} & l \in [0, 1) \\ \frac{w}{1-\delta} & l = 1 \end{cases} \tag{2}$$

With this definition at hand, it possible to express a HE as the solution to a system of functional equations.

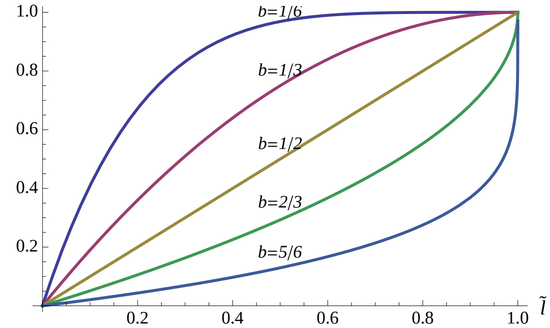
Proposition 2. *A HE exists if and only if there are functions* $(g(\cdot), v(\cdot))$ *such that for*

[13] More specifically, we require that the most senior member of a hierarchy is responsible for maintaining the public record, as they are the biggest beneficiaries, and have the lowest incentive to defect. A natural question that emerges is: what are the incentives that prevent agents from misrepresenting their status in the hierarchy? It is straightforward to show that agents have no incentive to misrepresent their status downwards, as this lowers their continuation value. More interesting is the incentive to misrepresent upwards. However, there is no incentive for upward misrepresentation either, since status is observable prior to matching and individuals are already trusted to the extent that their incentive constraints bind.

[14] Type $l = 1$ is already granted maximal trust by definition and therefore there is no need to worry about the agent's incentives if she were to be trusted 'more' off the equilibrium path.



(a) Density $g(\tilde{l}, 0)$



(b) Cumulative density $\int_0^{\tilde{l}} g(x, 0)dx$

Figure 1: **Advancement Probabilities as Seniority Rises, by Social Mobility**: The figure shows how the advancement probability and cumulative distribution changes at different levels of seniority $\tilde{l}$ for different levels of social mobility b. Notice that for $b = 1/2$, the graph is flat: advancement is independent of current seniority. For $b > 1/2$, there is *higher social mobility*: newcomers advance faster, and advancement slows as they become more senior. With $b < 1/2$, there is *limited social mobility*: newcomers advance more slowly, and advancement accelerates with seniority.

all seniority levels $l \in [0, 1]$, $v(\cdot)$ *satisfies* (2) *and the following conditions hold:*

$$l(s - w) \leq \delta(v(l) - v(0)), \qquad \textit{with equality if } l < 1; \qquad (IC_L^H)$$

$$l(s - w) \leq \frac{\delta^2}{N - 1}\left(w \int_l^1 g(\tilde{l}, l)\tilde{l}\ d\tilde{l} + c\right). \qquad (IC_R^H)$$

The left hand side of both (IC_L^H) and (IC_R^H) is the period gain for an agent trusted at seniority level l from cheating over cooperation. The (IC_L^H) constraint ensures agents prefer cooperation to strategies where they *cheat and then leave*, starting anew elsewhere. The (IC_R^H) rules out cheat and return defections. The right hand side is the expected punishment conditional on returning, which depends on the probability of both parties surviving and being re-matched, and an individuals' expected trust the next period.

Our focus on equilibria without any artificial withholding of trust is captured by the requirement that the (IC_L^H) binds for all types $l < 1$.

To allow us to solve for the equilibrium, we further restrict g to belong to the following family of advancement processes, parameterized by $b \in (0, 1)$:

$$g(\tilde{l}, l) = \frac{1 - b}{b} \times \frac{(1 - \tilde{l})^{\frac{1}{b} - 2}}{(1 - l)^{\frac{1}{b} - 1}}. \qquad (3)$$

This parametrization allows us to capture *social mobility* within the community in a straightforward way. Let $\mu_{\tilde{l}|l}$ denote the expected seniority $\tilde{l}$, given current seniority l:

$$\mu_{\tilde{l}|l} = \int_l^1 g(\tilde{l}, l)\tilde{l}\, d\tilde{l}$$

Observe that:

$$\mu_{\tilde{l}|l} = l + b(1 - l) \Leftrightarrow b = \frac{\mu_{\tilde{l}|l} - l}{1 - l}. \tag{4}$$

Although imposing a functional form for g is obviously restrictive, the form still allows for a rich family of advancement processes. The parameter b can be interpreted as the proportion of the remaining gap an individual faces between current seniority and the highest possible level of seniority that the individual expects to cover in one period, conditional on surviving. Recalling that l is normalized so that it indicates the level of trust in the HE, it is thus a gauge of *social mobility* (see Figure 1). When $b \to 1$, all individuals expect to get to the top in one period– i.e. there is extreme mobility, and individuals are trusted fully after their initial period. When, $b \to 0$ for all $\tilde{l} > l$, then $g(\tilde{l}, l) \to 0$, i.e. there is no mobility. An intermediate case is $b = 1/2$. In this situation, the probability of advancement for an individual of seniority l is uniform between l and 1, and thus each period, each individual expects to cover half the distance between their current seniority and the top rank of $l = 1$.

Characterizing the equilibrium reduces to showing that there is a b that satisfies the requirements of a Hierarchical Equilibrium.

Proposition 3. *A Hierarchical Equilibrium exists if and only if*

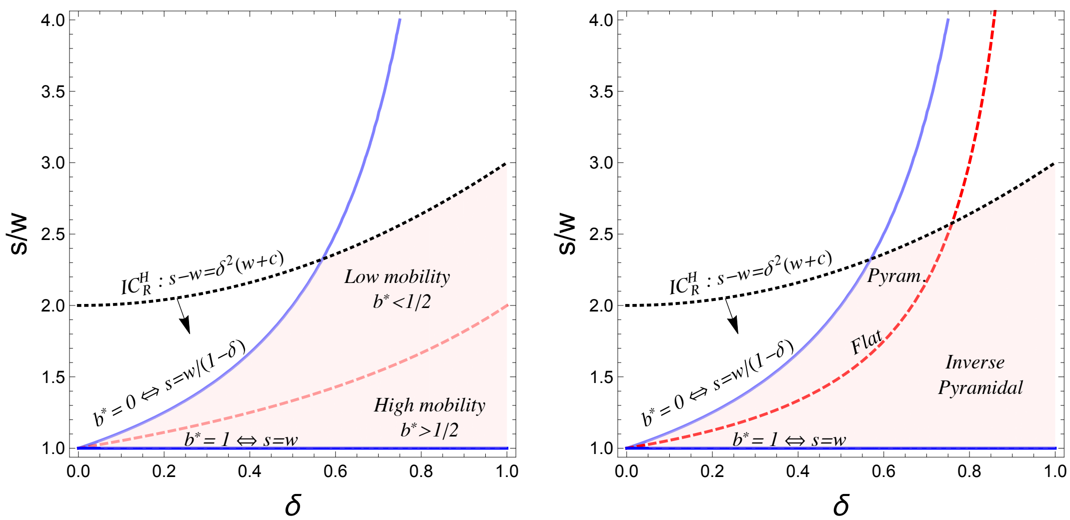
$$s < \frac{w}{1 - \delta}, \tag{5}$$

$$s - w \leq \frac{\delta^2}{N - 1}(w + c), \tag{6}$$

and is characterized by:

$$v^*(l) = \frac{w}{1 - \delta} - (1 - l)\frac{s - w}{s} \qquad \textit{and} \qquad b^* = 1 - \frac{s - w}{\delta s}. \tag{7}$$

First, notice that the cheat and return constraint (IC_R^H) is hardest to satisfy for the most senior $(l = 1)$ types. For them the ratio of current to future trust is highest (it equals one), and thus the temptation to defect today is greatest relative to the expected loss in future trust in the punishment phase. So, (6) is simply the (IC_R^H) computed for



(a) HE Existence region in the $(\delta, s/w)$ space assuming $c = 1$, $N = 2$ and shape of the equilibrium advancement process g^*.

(b) Same Existence region of subplot (a) but focusing on the shape of the induced steady state distribution f^*.

Figure 2

$l = 1$ which is basically the condition on group size of the identity investment benchmark. It follows that hierarchies do not make it easier to provide incentives to cooperate *within* groups.

Next, observe that, perhaps surprisingly, the existence regions for HE and the identity investment equilibrium actually are the same. Both require $s < w/(1-\delta)$, and therefore that shirking can be deterred with a maximal punishment of no trade in future periods. In the HE, when s is close to $w/(1-\delta)$, the hierarchical structure needs to have very low mobility ($b^* \to 0$), so that starting over in a new community is maximally unattractive ($v(0) \to 0$). Since the HE has a number of attractive features relative to the identity investment equilibrium— such as obviating the requirement that trust is denied to agents who would in fact be trustworthy— it is valuable to establish that these properties do not come at the expense of more restrictive conditions for existence.

Figure 2(a) depicts the existence region for a HE as a function of δ and s/w and relates them to the shape of the equilibrium advancement process. Recall that existence depends on four parameters, s/w, δ, N and c. In the figure, we set $c = 1$ and $N = 2$. The region enclosed by solid lines is defined by whether it is possible to find an advancement process (parameterized by b) to support a HE; thus, the upper and lower solid bounds represent the solution to (7) when $b^* = 0$ and $b^* = 1$, respectively. When s/w is such that s is close to the maximum long-term loss from deviating, $w/(1-\delta)$, then, as discussed above, $b^* = 0$. On the other hand, when s/w is close to 1 (so that it is also possible to satisfy $s < w/(1-\delta)$), there is no incentive to shirk and so cooperation can be sustained for all δ. Then, the HE has b^* close to 1 (in order to ensure that (IC_L^H) binds), so that newcomers advance to full trust very quickly.

The area below the dotted line is where the cheat and return constraint is satisfied. The dashed line in between solid lines in 2(a) is the locus where $b^* = 1/2$, that is, where the equilibrium is supported through a uniform equilibrium advancement process. It illustrates the boundary between *high (social) mobility* and *low (social) mobility* processes. Intuitively, the higher the incentive to shirk, the higher the patience required to sustain rapid advancement up the hierarchy.

We can also characterize the steady state density of the seniority level of individuals, denoted f^*, in this equilibrium.[15] This density is determined jointly by the survival rate δ and the advancement process (b^*), which also depends on δ. Since seniority is normalized to correspond to the level of trust, the shape of f^* describes the distribution of trustworthiness in the community. f^* solves the following functional equation for $l \in (0,1)$.

$$\overbrace{f^\star(l)(1-\delta)}^{\#\text{ individuals dying}} + \underbrace{f^\star(l)\delta\int_l^1 g^*(\tilde{l},l)d\tilde{l}}_{\#\text{ ind. promoted to upper levels}} = \overbrace{\delta\int_0^l g^*(\tilde{l},l)f(\tilde{l})d\tilde{l}}^{\#\text{ ind. promoted from lower levels}} + \underbrace{(1-\delta)g^*(0,l)}_{\#\text{ newcomers landing at level } l} \tag{8}$$

The above has exactly one solution given by:

$$f^*(l) = \frac{1-b^*}{b^*}(1-\delta)(1-l)^{\frac{1-b^*}{b^*}(1-\delta)-1}. \tag{9}$$

Taking the derivative with respect to l and rearranging, one can show that $f^*(l)$ decreases with seniority l if and only if:

$$\frac{s-w}{s} > \frac{\delta}{2-\delta}. \tag{10}$$

Observe that communities have *pyramidal* hierarchical structures—i.e., there are more individuals at the bottom than the top—when either the incentives to cheat are large ($s-w$) or if patience is low. Conversely, higher social mobility and more *egalitarian* inverse-

[15]More precisely, $\int_{l'}^{l''} f(l)dl$ equals the probability that an individual sampled at random at some arbitrarily large time t has seniority between l' and l'').

pyramidal communities are sustained in maximal trust hierarchies when the incentives to cheat and defect are weaker.

2(b) illustrates the existence region partitioned by the shape of the steady state distribution. Notice that the dashed line in Figure 2b (representing a uniform distribution over seniority) would fall to the left of the dashed line if superimposed on Figure 2a (representing a uniform advancement process), and the wedge between them increases with δ. Intuitively, when individuals are long-lived, even with low social mobility, they will mass at the top of the distribution. So an equilibrium advancement process with low social mobility does not imply a pyramidal hierarchy in the steady state. However, a pyramidal steady state distribution does imply the presence of low social mobility.

We can also examine aggregate welfare. The closed form expression for aggregate average per period surplus is:

$$(1-\delta)\int_0^1 \tilde{v}(\tilde{l})f^*(\tilde{l})d\tilde{l} = \frac{(w-(1-\delta)s)((1-\delta)s+2\delta w - w)}{\delta^2 w}$$

Naturally, average per period surplus decreases with temptation to cheat $s - w$ and increases with patience δ, Further, average surplus approaches $\frac{w}{1-\delta}$ whenever s approaches w and approaches 0 whenever s approaches $w/(1-\delta)$.

One may wonder to what extent (A1)-(A3) constrain the equilibrium set. As discussed above, the partial order (A1), together with monotonic strategy profiles, defines what we mean by a 'hierarchical institution.' (A2) restricts attention to hierarchical structures to those where there is upward social mobility. Notice that, when payoffs are increasing in seniority, what matters for incentives is whether G is such that the *expected* period payoff in $t+1$ increases with an seniority at time t. Thus, allowing for *demotions* should not change the nature of the incentive constraints provided that individuals are still promoted *on average* so that their expected payoff goes up. While we do not provide a formal treatment of this case, it is intuitive that one can relax (A2) and construct equilibria similar in spirit.[16] The assumption that G is smooth is used in the characterization result (Proposition 2). It guarantees that the functional operator defined by equation (2) is well-defined (G has a density function) and that the operator preserves continuity. This then implies— by the Contraction Mapping Theorem—that $v^*(l)$ is continuous.[17]. The

[16]Note also we allow individuals to *fall from grace*– descend the hierarchy– in a limited sense, by allowing them to start anew in a different community.

[17]In equation (2), l is basically a Markovian shock over which the expectation is taken. For this operator to preserve continuity it is sufficient that the transition function defining the stochastic Markovian process satisfies the Feller property, a much weaker requirement. See Stokey and Lucas, section 8.1

functional form restriction (3), although still sufficiently flexible to capture the spectrum of relative mobility, is made for tractability. Crucially, it allows us to recover the closed form characterization presented in Proposition 3.

Hierarchies with Trading Selectivity

Up until now we assumed that having chosen a community, individuals are randomly matched with other members. We now consider a setting capturing the idea that individuals can choose whom to interact with, even within a community. The idea is that if some community members are more trustworthy, they would naturally become the focus of trust. That is, they will be chosen by more than one principal and obtain on average more than one trade per period in equilibrium. One extreme example is a community in which, say, the highest in rank works as an agent for everyone else while the lowest ranked community members do not work at all.

Though conceptually straightforward, permitting trading selectivity within hierarchies raises a number of modeling challenges when considering large groups. To accommodate the more complex possibilities of selective interaction in such groups, we introduce exogenous *capacity constraints*, that also allow us to think about technological change. Notice that we already have been implicitly assuming that such capacity constraints exist in the previous section: each player was limited to be trusted at most once. We now relax those capacity constraints, assuming instead that each individual can perform at most $J \geq 1$ tasks. In other words, individuals can be entrusted as an agent by more than one principal.

In what follows, we keep the total number of potential trades in the community the same (equal to N), but allow the distribution of trades to change. As pointed out already, if some individuals, (say the highest ranked ones) are *more* trustworthy then all individuals would want to select them. Typically, these highest ranked agents will not be able to serve all members of a community. To capture trade selectivity in this context, we make the natural assumption that individuals randomize only when indifferent: they expect to be randomly assigned to an agent among those that are most trustworthy and that all agents work at capacity. This implies that we are looking for efficient equilibria with stable matches.[18]

[18] I.e. where there is no principal and agent pair that would rather be matched to someone other than their equilibrium partner and trade is maximized.

A key challenge is that, if higher ranked individuals work relatively more, individuals' incentives depend on their *relative* seniority and thus the entire realization of seniorities every period. As an illustration, take the simple example in which agents can perform up to $J = 2$ tasks each per period. In this world, it is possible for half of the individuals to work for everyone else. Suppose that period $t + 1$ trades are concentrated among those who happen to be in the upper half of the hierarchy in that period. What matters for incentives at time t, among other things, is the subjective probability of being in the upper half of the hierarchy upon returning at time $t+1$. Clearly, such probability depends on the entire realization of seniorities at time t which basically acts as an N-dimensional state variable.

We can make progress, however, by focusing on two limit cases that are solvable either analytically or numerically. First, we look at the case where interactions are very valuable (w large) relative to the short term gain from shirking $s - w$. As we shall see, in this case, groups can be arbitrarily large in equilibrium. This simplifies the problem because the empirical distribution of seniority levels converges to the steady state distribution. So, it does not change over time or across communities.[19] This allows to solve the model analytically and hence, to derive a number of results. To show that our findings do not depend on groups being large, we further solve numerically the polar opposite case: of groups with three individuals (see appendix B). This analysis shows that the key insights, particularly with respect to robustness in the presence of other institutions, are still valid.

Formally, the hierarchical setting with trading selectivity is the same as the hierarchical setting with random matching, but with the following modifications. While each individual continues to have one task they would like entrust an agent with, we now allow individuals to act as an agent for up to J principals, where, as discussed above, J can be thought of as an exogenous *capacity constraint* brought upon by technological change. So, in each period, as few as N/J individuals could be active as agents and thus gain all the period surplus. For simplicity, assume N/J is integer and $J \leq N - 1$.[20] As in the previous section, individual seniority evolves over time according to (3).

Let $\boldsymbol{\ell}$ denote a vector whose i-th component l_i is the seniority of individual i. Also let $\hat{l}(\boldsymbol{\ell})$ denote the threshold seniority level such that J multiplied by the number of agents

[19] In the example above this means that there will be a (time-invariant) threshold seniority that separates the upper half from the bottom half of the population. So, the probability that i is in the upper half in $t + 1$ upon returning depends only on i's seniority at time t.

[20] An individual cannot work for herself. So we set aside the case $N = J$ as it requires additional technical assumptions that would blur the exposition.

with seniority greater or equal to $\hat{l}(\boldsymbol{\ell})$ equals the total number of tasks N.[21] In other words, $\hat{l}(\boldsymbol{\ell})$ is such that assigning J principals and thus J tasks to each individual with seniority weakly higher than $\hat{l}(\boldsymbol{\ell})$ would exhaust all N trades. Clearly, if $J = 1$ then $\hat{l}(\boldsymbol{\ell})$ is the smallest element of $\boldsymbol{\ell}$. If $J = N - 1$ then $\hat{l}(\boldsymbol{\ell})$ is the largest element of $\boldsymbol{\ell}$.

A maximal *Hierarchical Equilibrium with Trade Selectivity* (HE-TS) is the same as a HE with the following amendments. On the equilibrium path, given an arbitrary realization of seniorities $\boldsymbol{\ell}$, all principals are randomly matched to an agent with seniority greater or equal than $\hat{l}(\boldsymbol{\ell})$. The randomization process makes sure that no individual is matched as a principal to more than J agents. So, on the candidate equilibrium path, $\hat{l}(\boldsymbol{\ell})$ separates 'active' from 'idle' individuals. l types are trusted at scale l and work. Conditional on the agent returning in the period after cheating, the punishment phase is a one period Nash reversion to no trust/shirk within the relevant bilateral trading relationship. In the continuation game after any one-stage deviation to the above, the principal(s) always returns and the agent always leaves.

At time t, i's subjective probability to belong to the active subset of the population in $t + 1$ depends on her own seniority l_i and on other members' seniorities, denoted $\boldsymbol{\ell}_{-i}$, as does the continuation value on the candidate equilibrium path (which we will denote $v(l_i, \boldsymbol{\ell}_{-i})$). If a Hierarchical Equilibrium with Trade Selectivity exists, an individual's payoff along the equilibrium path is Jw if $l_i \geq \hat{l}(\boldsymbol{\ell})$ and is zero otherwise. Thus we can express $v(l_i, \boldsymbol{\ell}_{-i})$ as follows:

$$v(l_i, \boldsymbol{\ell}_{-i}) = \begin{cases} Jw \int_{\tilde{\boldsymbol{\ell}}} \left(\tilde{l}_i \mathbb{I}_i(\tilde{\boldsymbol{\ell}}) + \delta v(\tilde{l}_i, \tilde{\boldsymbol{\ell}}_{-i}) \right) \mathcal{G}(\tilde{\boldsymbol{\ell}}, \boldsymbol{\ell}) \, d\tilde{\boldsymbol{\ell}} & l_i < 1 \\ \frac{Jw}{1-\delta} & l_i = 1 \end{cases}, \tag{11}$$

where $\mathcal{G}(\tilde{\boldsymbol{\ell}}, \boldsymbol{\ell}) := g(\tilde{l}_1, l_1) \times ... \times g(\tilde{l}_N, l_N)$ denotes the joint density over future seniorities $\tilde{\boldsymbol{\ell}}$ given current seniorities $\boldsymbol{\ell}$, and $\mathbb{I}_i(\boldsymbol{\ell})$ is an indicator function equal to 1 if $l_i \geq \hat{l}(\boldsymbol{\ell})$.

We now look for necessary and sufficient conditions for the above candidate equilibrium to exist. First, consider the incentives to cheat and return. Within each bilateral trading relationship, the decision to work is incentive compatible if and only if the expected discounted surplus loss due to the one period reversion to shirk / no trust between the agent and its current partner is low enough. Such loss depends on the probability of meeting a cheated trading partner again in $t + 1$ and on the expected surplus lost,

[21] Notice that $\hat{l}(\cdot)$ is well defined for all $\boldsymbol{\ell} \in [0, 1]^N$ such that $l_i \neq l_j$ for all i, j with $i \neq j$. This is not an issue in what follows since $\boldsymbol{\ell}$ is drawn at the beginning of every period from a continuous distribution over the unit square and $l_i = l_j$ is a zero probability event.

conditional on meeting that agent. If the agent is active at time $t+1$ the probability of being re-matched to the same principal, conditional on both surviving is $\frac{J}{N-1}$. As argued in the previous section, this constraint binds first for the highest seniority ($l=1$) type. So, the expected discounted surplus loss is $w+c$ times $\delta^2\frac{J}{N-1}$. Cheating and returning is thus not profitable if and only if:[22]

$$s - w \leq \delta^2 \frac{J}{N-1}(w+c).$$

Agents prefer working and returning to cheating and leaving if and only if the opportunity cost of starting afresh somewhere else is lower than the present gain of cheating. If the agent is bound to leave, then in equilibrium it is optimal to cheat simultaneously against all trading partners. So, working is preferred to cheating and leaving if and only if the short term gain of shirking simultaneously against all J trading partners is low enough:

$$Jl_i(s-w) \leq \delta(v(l_i, \boldsymbol{\ell}_{-i}) - v(0)).$$

The following proposition brings this analysis together.

Proposition 4. *A HE-TS exists if and only if there exists a tuple* $(b, v(\cdot))$ *such that for all seniority profiles* $\boldsymbol{\ell} \in [0,1]^N$, $v(\cdot)$ *satisfies* (11) *and the following conditions hold:*

$$Jl_i(s-w) \leq \delta(v(l_i, \boldsymbol{\ell}_{-i}) - v(0)), \qquad \text{with equality if } \hat{l}(\boldsymbol{\ell}) < l < 1; \qquad (IC_L^{TS})$$

$$s - w \leq \delta^2 \frac{J}{N-1}(w+c). \qquad (IC_R^{TS})$$

Notice that when each agent can perform at most one task ($J=1$) all individuals work: $\hat{l}_t = 0$ and we are back to the no trade selectivity case analyzed in the previous section.Recall that i's period payoff on the equilibrium path depends only on whether l_i exceeds the threshold $\hat{l}(\boldsymbol{\ell})$. When $(s-w)/w$ goes to zero, the incentive compatibility constraint (IC_R^{TS}) is maximally relaxed and group size N such that it holds with equality goes to infinity. This implies that:

(i.) the empirical distribution of seniority levels $\boldsymbol{\ell}$ approaches the steady state distribution

[22] when $J>1$, agents may cheat simultaneously on multiple principals. However, as players cannot condition on what happens within other trade relationship and the probability of either trade partner surviving is independent on others', incentives within bilateral relationships are not affected by actions taken outside the relationship.

$f(l)$:

$$f(l) = \frac{1-b}{b}(1-\delta)(1-l)^{\frac{1-b}{b}(1-\delta)-1}, \tag{12}$$

(ii.) the threshold $\hat{l}(\boldsymbol{\ell})$ approaches a constant defined by:

$$\bar{l} := \{l : J\int_l^1 f(l')dl' = 1\}. \tag{13}$$

In turn, this implies that the continuation payoff $v(\cdot)$, is also constant in $\boldsymbol{\ell}_{-i}$ reducing the dimensionality of the problem and allowing us to characterize the equilibrium. With a slight abuse of notation we drop the second argument of v in what follows.

Proposition 5. *In the limit as $\frac{s-w}{w} \to 0$, a HE-TS equilibrium exists with:*

$$v^*(l) = \begin{cases} \frac{(1-l)^{-\alpha}}{(1-\bar{l})^{-\alpha}}\left(J\frac{w}{1-\delta} - J(1-\bar{l})\frac{s-w}{\delta}\right) & 0 \le l \le \bar{l} \\ \\ J\frac{w}{1-\delta} - J(1-l)\frac{s-w}{\delta} & \bar{l} \le l \le 1 \end{cases} \tag{14}$$

$$b^* = 1 - \frac{s-w}{\delta s}, \tag{15}$$

where $\alpha := \frac{1-b^*}{b^*}(1-\delta) < 1$.

Let us interpret how the equilibrium works. Trading selectivity allows for further redistributing of period surplus from lower ranked individuals to higher ranked individuals. This makes it more costly for higher ranked individuals to cheat and start anew in a different community. In turn this creates the incentive for others to select and trust a higher ranked individual.

Notice further that by focusing trades among individuals with higher levels of seniority, hierarchies with trade selectivity also increase the fraction of the period surplus that is appropriated by the returning members of the community. This allows for two key robustness features of hierarchies relative to other institutions.

First, *hierarchies can exist in the presence of other groups that follow different strategies, including, in the extreme, a community that lowers its barriers or removes its institutions completely, and trusts newcomers fully and from the start.* This is because if social mobility is high enough, the sharing of the future surplus that accrues to returning individuals if they rise in status can more than compensate for the losses that individuals with low seniority incur in hierarchies due to the fact that they may experience an initial period without being selected for trade. Formally, all individuals find it more attractive to start afresh in hierarchical community than in a hypothetical institution-free community with full cooperation if $v(0)$ is larger than $w/(1-\delta)$. That is if:

$$J\frac{w}{1-\delta} - J\frac{s-w}{\delta} \geq \frac{w}{1-\delta}. \tag{IF}$$

(IF) can be satisfied only if $J > 1$. Although it may seem surprising that it is possible for an agent to expect, on average, more than the full per-individual surplus (discounted to factor in survival), our model has the flavor of an overlapping generations model. Each individual expects other members of the community to die with positive probability, and when they do, their replacements will on average have lower seniority. (IF) also implies that cheating within a hierarchy and then leaving for an institution free community with full cooperation cannot be optimal.

Second, *hierarchical communities with sufficient social mobility can in fact, lead to the breakdown of cooperation in institution-free communities.* Individuals belonging to a hypothetical institution free community with full cooperation find it profitable to cheat and leave for the hierarchy if and only if

$$s + \delta v(0) \geq w + \delta\frac{w}{1-\delta} \tag{16}$$

Since $s > w$, this condition is implied by (IF).

Thus, hierarchies can not only be robust to other institutions, but their presence can also make it impossible to sustain alternative institutional arrangements, when agents have the option to cheat and leave.

As a robustness check, in the online appendix we solve the model for the polar opposite limit case: where parameter values are such that the equilibrium size of communities is the smallest $N = 3$. Using numerical methods, we show that a trade selectivity hierarchical equilibrium exists and, in addition, can satisfy the robustness conditions (IF) and (16).[23]

[23] The only difference is that individuals may leave on the equilibrium path if seniority realizations are such that some individuals are left behind and thus find it optimal to leave and start afresh rather than play catch up.

4 Applications

Our model may be applied to shed new light not only on the design and robustness of institutions that can sustain cooperation across a range of challenging contracting environments, but also on a number of puzzles related to early human economic and political development. In the Online Appendix, we describe several applications in detail; here, we highlight the main points of three of these applications before analyzing a fourth in greater depth.

A first application of our model is in explaining the emergence of hierarchy in early human societies. A literature on this emergence focuses on the technological change from pastoralism to agriculture, the advantage of first movers in acquiring geographically desirable locations or the storability of crops in allowing coercion (eg Dow and Reed, 2013, Boix, 2015, Mayshar, Moav, Neeman, and Pascali, 2015). Our model suggests an alternative, though potentially complementary, explanation: as population density increases, "cheat and leave" defections become easier, and social hierarchy provides a way to sustain cooperation in the face of this. In the Appendix, we use data from Binford's 2001 study of 339 non-agrarian hunter-gatherer societies, described at first contact with anthropologists. We define *egalitarians* to include all societies classified as "generic hunter-gatherers" or "generic hunter-gatherers with institutionalized leaders," while *hierarchical* societies include all societies classified as "wealth-differentiated hunter-gatherers" or "stratified or characterized by elite and privileged leaders." We show that as population density increases, hierarchical societies become much more prevalent than egalitarian ones.

Another implication of our model is that hierarchy can help sustain trust in larger groups, so long as seniority confers higher levels of trust. As long as there is common knowledge of seniority, it is not necessary for individuals to keep track of individual trading histories. The example of the ascendance of the Roman Republic over the ethnically-delimited organization among the city-states of Carthage and Greece illustrates how seniority-based systems can achieve greater scale and trade. Citizenship in "egalitarian" Athens, for example, derived from ethnic descent from one of four tribes. While citizens had special privileges, "foreigners" were excluded from economic opportunities. Rome, in contrast, allowed many different ethnic groups to become members. Roman society was hierarchical, with ranks including slaves, citizens, junior and senior office holders, and senators. There was mobility across ranks, including for slaves, so much so that by the second century CE, most citizens had slave ancestry. While both contemporaries and modern historians have contended that the growth of the Greek city-states was constrained by their institutions, Rome's hierarchy arguably sustained trust and cooperation at very large scale.

A third example concerns the ability to sustain trust in online communities, such as open source software development and online forums. Many such communities have hierarchies; for example, more senior members have voting rights and help determine the direction of projects. Online forums typically note some form of seniority for experienced members. As in our model, more senior members are trusted more, and as such have greater incentives to return, which in turn makes them more trustworthy.

Pastoralist and Agriculturalist Hausa in Nigeria The case of the Hausa in Nigeria further illustrates how emergent hierarchies sustain trust in a modern society with weak legal institutions while undermining cooperation in nearby egalitarian communities.[24]

Among the Hausa, it is religious identity– particularly Islam– that plays an important role in shaping access to commercial opportunities. However, Hausa Muslims and non-Muslims (known as Maguzawa) have common ethnic origins and have lived intermixed historically, with Maguzawa seen by Muslims as being the "original" Hausa (Last, 1999).[25] Though Maguzawa now constitute fewer than 2% of the 50 million modern Hausa-speakers, mass conversion to Islam is relatively recent: in 1900, only 5% of contemporary Hausaland were Muslim (Salamone, 2010).

Yet, despite their similarities, and the ability of individuals to move relatively freely between Muslim and Maguzawa groups within the Hausa, there remain remarkable differences in their social organization and levels of trust. The Muslim Hausa have a well-developed hierarchical structure that transcends lineage, while the Maguzawa have well-established norms of egalitarianism (King, 2006, Last, 1979, 1999, Barkow, 1974, Salamone, 2010). As in the hierarchical equilibrium, access to enhanced trading opportunities is directly related to the extent to which the agent is embedded in a network of trust-based relations. Those with the highest status include long-distance Muslim traders, who traded along pilgrimage routes across the Sahara (King, 2006, Barkow, 1974, Cohen, 1969). Those involved in smaller-scale commerce or professions involving

[24] Of the 20 sub-Saharan African states surveyed by Afrobarometer in 2008, Nigerians ranked last, with only 9.50% stating that they trusted non-relatives they knew "a lot", compared to 25.50% of other sub-Saharan African respondents. Formal third party enforcement is also perceived to be among the weakest in sub-Saharan Africa.

[25] As Last (1979)[239] notes, the Maguzawa were "considered one of "us", therefore not targets for jihad even in the early nineteenth century ... There was apparently, and still is, no urge among the Muslims to convert the non-Muslims among their neighbors and subjects ... "

less trust also have lower perceived status (Barkow, 1974). Qualitative accounts stress how among the Hausa, "a good deal of business is conducted with handshakes and one's word" (Salamone, 2010)[p.3].

In contrast to the Muslim Hausa, the Maguzawa are a good example of an 'institution-free' competing group. Maguzawa have strong norms of egalitarianism and reciprocal obligation (King, 2006, Last, 1979, 1999, Barkow, 1973, Salamone, 2010). They are also perceived as being completely open and welcoming to new entrants (Barkow, 1973, Last, 1999). Though, the Maguzawa have lived in close proximity to the Muslims for centuries, there is no ambiguity in classification: both groups maintain clear (though reversible) visual symbols of distinction, including different styles of clothing and restrictions on women (Barkow, 1973, Last, 1999).

Hausa thus have had the option of choosing to be Muslim and thus joining the hierarchical structure of the Muslim Hausa, or choosing to remain or even switch to being Maguzawa (Last, 1979).[26] Consistent with the model, new converts to Islam are *actually poorer* than non-Muslim Hausa (Last, 1979, 1999, Barkow, 1974).[27] Yet, despite the drop in wealth, it appears that it is the possibility of advancement in the Muslim hierarchy (and thus of being the focus of coordinated trades and thus increased wealth in the future), that drives individuals to convert.[28]

Further, despite common ethnic, linguistic and geographic endowments, and the relative lack of barriers to conversion to Islam (Last, 1979), the differences in the degree of trust between Hausa religions are remarkable. Table 1 shows the proportion of Nigerian respondents to the Afrobarometer survey who reported their willingness to trust others from their country "a lot" or "somewhat". As the table suggests, while both non-Muslim and Muslim Hausa tend to report a greater willingness to trust others in general compared to non-Hausa Nigerians, there are remarkable differences within the Hausa according to the degree they trust people within and outside their own families. While a greater proportion of non-Muslim Hausa show a willingness to trust relatives, Muslim Hausa are more likely to report that they trust both non-relatives that they know and strangers more generally.

Consistent with our model, rather than the 'institution-free' community undermin-

[26]In fact, consistent with the model, new Muslim converts who "fail" in trade can and do change their identity to Maguzawa (Last, 1999).

[27]Ironically, in the early years of independence, due to their relatively high wealth, the agriculturalist Maguzawa were classed as "merchants" and taxed at a higher rate on average (Last, 1999).

[28]Last (1979) describes: "Though it is clear to non-Muslims that the Muslim is often poorer and lives a more constricted social life, the prospect is of wealth through trade . . . (239)"

		Non-Muslims	Muslims	All
Proportion claiming to trust:				
Non-Hausa	Relatives	0.678	0.741	0.692
	Others they know	0.386	0.464	0.403
	Other Nigerians	0.247	0.289	0.257
	Observations	1,394	401	1,795
Hausa	Relatives	0.918	0.852	0.858
	Others they know	0.469	0.635	0.620
	Other Nigerians	0.367	0.469	0.459
	Observations	49	480	529

Table 1: **Trust in Nigeria by ethnicity and religion**
Proportion of population that trust other Nigerians "somewhat" or "a lot" (as opposed to "not at all" or "just a little".) source: Afrobarometer 2008-2009

ing cooperation among the Muslim Hausa, the Muslim community has attracted more converts, despite the fact that converts tend to be poorer on average, and it has been cooperation among the Maguzawa that has largely broken down. Once living in close proximity to Muslims in the towns of Hausaland, the remaining Maguzawa have become specialized economically in largely autarkic farming activities, and have become more geographically dispersed, both from Muslims and from each other (Greenberg, 1947, Last, 1979, 1999).[29]

5 Conclusions

This paper has analyzed the endogenous formation of groups that enable their members to sustain trust in environments where there are many possible partners, outside options are strong and legal enforcement is not available. An important set of "identity investment" institutions sustain loyalty to a group and cooperation within groups by creating barriers to starting anew. Seniority structures, however, can support loyalty and within-group cooperation without requiring artificial barriers to entry or withholding of trade. In fact, in environments where individuals can choose their trading partners, seniority structures can be designed that are robust to the entry of institution-free groups that would lead to the failure of trust in groups that are sustained by entry barriers. Further, such seniority structures can actually undermine cooperation in institution-free groups.

[29]In the 1970s, virtually all Hausa in areas with population densities above 200 per square mile had become Muslim (Last, 1979).

The resilience of seniority- based cooperative hierarchies and the corresponding decline of trust and cooperation in institution-free groups appears to mimic a number of environments where new venues for exchange have emerged, from open source software communities to early human societies. But it also hints at the dynamics of spontaneous order to expect in both historical and contemporary settings where no institutions exist. Though the formation of entry barriers based upon "cultural" distinction can sustain cooperation among sub-groups, these will be undermined by the continued presence of institution-free groups. Social hierarchies, in contrast, may survive. In the absence of third-party enforcement, hierarchies may thus emerge as a common early organizational form in new venues for trade.

We are also able to explain an important paradox, how absent coercive power supportive of secure property rights, that small inequalities that underpin the historic emergence of elites in human societies might persist and grow rather than being simply expropriated in societies with strong egalitarian norms (Sterelny, 2013, Bowles and Choi, 2013). In our hierarchical equilibria, it is increased bilateral trust that provides the wealth associated with status, and being intangible, such trust cannot be expropriated. Further, unlike social networks which are often person-specific, social hierarchies are "impersonal". Individuals can trade with and trust others that they have never met, with rank providing sufficient information for trust, and the patterns of cooperation in a social hierarchy can survive beyond the individuals that populate them. Rank indicates that individuals benefit from remaining in a community, and thus are more trustworthy.

A Appendix: Formal Proofs

A.1 Proof of Proposition 1

(IC_L) and(IC_R) are definitionally necessary for a HE to exist. If (IC_L) were violated, an agent would have an incentive to cheat and leave so it wouldn't be optimal to trust him in the first place. Similarly, if (IC_R) were violated, an agent would be better off cheating and returning.

We now show that conditions (IC_L) and (IC_R) are also sufficient. On the equilibrium path (IC_L) and (IC_R) imply that, no individual has an incentive to deviate given the candidate equilibria of the continuation game following a deviation. Finally, we need to show that off the equilibrium path the principal always returns so that he would be there to provide the punishment if the agent were to return. In a candidate equilibria where both individuals return, they must have the same continuation payoff. It follows that in order to support an identity investment equilibrium one can rely on strategies where principals always return and an agent leaves in the period after cheating if and only if the continuation payoff from returning to a community with one enemy around is lower than that of starting afresh. Finally, notice that there is no profitable deviation for an agent to defect for a few periods and then leave, since the continuation payoff in the period after the punishment took place is equal to the continuation payoff on the equilibrium path.

A.2 Proof of Proposition 2

(IC_L^H) and (IC_R^H) are definitionally necessary for a HE to exist. We now show that they are also sufficient. On the equilibrium path (IC_L^H) and (IC_R^H) imply that, no individual has an incentive to deviate. Next, we now show that off the equilibrium path there must be an equilibrium of the continuation game in which the agent leaves and the principal returns to guarantee a punishment. The principal expects the agent to leave so returning is optimal as $v(l) > V(0)$. The agent optimally leaves after cheating if $v(0)$ is weakly larger than the continuation payoff from returning in a community with one enemy around. Such payoff is highest for $l = 1$ agents. Continuity of v implies that the (IC_L^H) constraint holds with equality for type $l = 1$. So all types are indifferent between working and returning and cheating and leaving while weakly preferring to work to cheating and returning (strictly so, for all $l < 1$). It follows that in order to support a HE, one can rely on strategies where agents always leave and principals always return in the period after cheating. Given this, the argument proceeds along lines identical to those of Proposition 1, and is therefore omitted.

A.3 Proof of Proposition 3

If an equilibrium exists then its associated value function must satisfy (2) (reported here for convenience):

$$v(l) = \begin{cases} \int_l^1 g(\tilde{l}, l)(\tilde{l}w + \delta v(\tilde{l}))d\tilde{l} & l \in [0, 1) \\ \frac{w}{1-\delta} & l = 1 \end{cases}, \tag{A1}$$

We now show that there is only one one family of functions that solves the above problem parametrized by b. First, notice that the set of period payoffs on and off the equilibrium path is bounded from above by s and from below by 0. $v(l)$ is a present value of a stream of bounded payoffs and is thus bounded from above by $\frac{s}{1-\delta}$ and bounded from below by 0. It follows that $v(\cdot)$ must be bounded. (A1) defines an operator T that maps the set of bounded functions into itself. T satisfies the monotonicity and discounting conditions of Blackwell's theorem and therefore is a contraction with modulus δ. By the contraction mapping theorem there exists a unique fixed point. Since the set of bounded continuous functions is closed and the operator preserves continuity, it follows that the unique fixed point is bounded and continuous. Furthermore, by the second theorem of calculus, T maps continuous functions into differentiable functions, thus the unique fixed point is differentiable.

Differentiating both sides of the functional equation (A1) leads to a linear ordinary differential equation which can be solved up to constant c_1.

$$v(l)(c_1) = c_1(l-1)^{-\frac{1-b}{b}(1-\delta)} + \frac{w}{1-\delta}\left(\frac{1 + \frac{b}{1-b}(1-\delta)l}{1 + \frac{b}{1-b}(1-\delta)}\right). \tag{A2}$$

The constant is then pinned down by the boundary condition $v(1) = w/(1-\delta) \Rightarrow c_1 = 0$. The fixed point is:

$$v(l) = \frac{w}{1-\delta}\left(\frac{1 + \frac{b}{1-b}(1-\delta)l}{1 + \frac{b}{1-b}(1-\delta)}\right). \tag{A3}$$

Plugging (A3) in the (IC_L^H) constraint pins down $b^\star$:

$$l(s-w) = \delta(v(l) - v(0)) \Leftrightarrow b^\star = 1 - \frac{s-w}{\delta s}. \tag{A4}$$

Plugging $b^\star$ and in (A3) leads to

$$v^*(l) = \frac{w}{1-\delta} - (1-l)\frac{s-w}{\delta}, \tag{A5}$$

Finally $N^\star$ solves (IC_R) with equality for $l = 1$:

$$(s-w) = \frac{\delta^2}{N-1} w \Leftrightarrow N^* = \frac{\delta^2 w}{s-w} + 1. \tag{A6}$$

Notice that $b^\star > 0$ if and only if $s < \frac{w}{1-\delta}$. So an equilibrium exists if an only if this condition is satisfied.

A.4 Proof of Proposition 4

$(IC_L^{TS})$and (IC_R^{TS}) are definitionally necessary for a HE to exist. We now show that they are also sufficient. If a player is bound to leave then it would simultaneously cheat against all its current trading partners. So, on the equilibrium path, (IC_L^{TS}) guarantee that cheating and leaving cannot be optimal. Next consider the incentives to cheat and return to the home market, conditional on principals returning. As survival processes are independent, the expected discounted value of the trade lost due to the reversion to one-period punishment within a bilateral relationship does not depend on what happens outside that relationship. It follows that the agent either cheats against all of its trading partners and returns or works (assuming he works if indifferent). (IC_R^{TS}) guarantees that working is always optimal.

Given this, the argument proceeds along lines identical to those of Proposition 2, and is therefore omitted.

A.5 Proof of Proposition 5

If an equilibrium exists, the payoff on the equilibrium path is Jlw if $l \geq \bar{l}$ and zero otherwise. Let $\mathbb{I}(l_i \geq \bar{l})$ denote a function equal to one if the argument is true and zero otherwise. As the (period) payoff and $\bar{l}$ do not depend on $\boldsymbol{\ell}_{-i}$, the present value of the payoff stream $v(\cdot)$ is necessarily constant over $\boldsymbol{\ell}_{-i}$.

Consider the functional equation (11) defining the equilibrium value function. It defines an operator T that maps the set of bounded functions into itself. T preserves continuity and satisfies the monotonicity and discounting conditions of Blackwell's theorem so its unique fixed point is bounded and continuous. Since the set of bounded continuous functions is closed and the operator preserves continuity, it follows that the unique fixed point is bounded and continuous. Furthermore, by the second theorem of calculus, T maps continuous functions into differentiable functions, thus the unique fixed point is differentiable.

Integrating out $\boldsymbol{\ell}_{-i}$ leads to equation (dropping the argument $\boldsymbol{\ell}_{-i}$):

$$v(l_i) = \begin{cases} \int_{l_i}^{\bar{l}} g(\tilde{l}_i, l_i)\delta v(\tilde{l}_i)d\tilde{l}_i + \int_{\bar{l}}^{1} g(\tilde{l}_i, l_i)\left(J\tilde{l}_i w + \delta v(\tilde{l}_i)\right)\, d\tilde{l}_i & l_i \in [0, \bar{l}) \\ \int_{l_i}^{1} g(\tilde{l}_i, l_i)\left(J\tilde{l}_i w + \delta v(\tilde{l}_i)\right)\, d\tilde{l}_i & l_i \in [\bar{l}, 1) \end{cases}, \tag{A7}$$

Differentiating (A7) with respect to l_i leads to two linear ordinary differential equations (ODEs) which can be solved up to constant. Consider first $l \geq \bar{l}$. using the boundary condition $v(1) = \frac{Jw}{1-\delta}$ allows to find an exact solution up to parameter b . Plugging (11) in the (IC_L^{TS}) constraint pins down $b^{\star}$. Consider now $l \leq \bar{l}$. imposing continuity at $\bar{l}$ give a boundary condition $v(\bar{l})$ allowing to find an exact solution to the corresponding ODE.

B Online Appendix: Trading Selectivity with $N = 3$

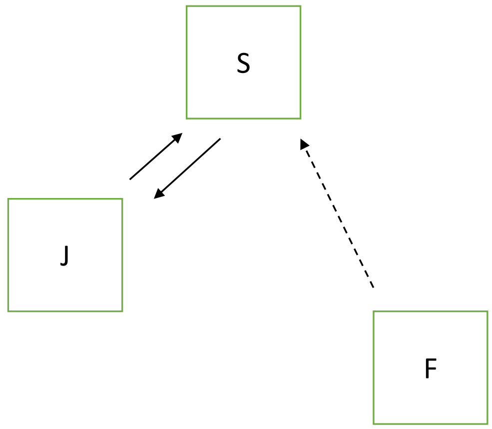


Figure 3: **Example: One-sided trust relationships in a three person hierarchy with trading selectivity before a newcomer (F) decides to leave or stay.**

We focus on studying the simplest hierarchy –one with three individuals– and rely on numerical methods to characterize the key equilibrium objects.

For convenience, given an arbitrary vector of seniorities $\boldsymbol{\ell} = (l_1, l_2, l_3)$ (from now on referred to as the 'state'), we call those individuals endowed with the highest, second-highest and lowest seniority as (S)eniors, (J)uniors and (F)reshmen, respectively.[1] Notice that a community's seniority rank ordering can differ every period: while there is a Senior every period, her identity (i.e. index) can change depending on the realization of the advancement process, the survival of other individuals and their choices to stay in the community. For example, an individual can stop being senior if either of the other two members draw a higher seniority.

B.1 Equilibrium notion

In order to further simplify the analysis we characterize equilibria where individuals are either trusted at full scale or not trusted. Specifically, we look for equilibria where (i.) On

[1]This involves a slight abuse of notation as if one or more individuals have the same seniority (a probability zero event), then we assume that seniority is randomly assigned. So $l_i = l_S$ is, strictly speaking, necessary but not sufficient for i to be senior.

the equilibrium path J trusts S at full scale and returns. (ii.) S trusts J at full scale and returns. F either fully trusts S and returns or trusts no-one and leaves. The punishment phase triggered by a deviation to the above in $t-1$ or in t is a one period Nash reversion to no trust/shirk in period t within the relevant parties. In the continuation game after any one stage deviation to the above, the principal(s) always returns and the agent always leaves.

Allowing for the possibility that cooperation breaks down between F and S is a key difference with the HE setting. F leaves if and only if the state is such that the continuation payoff from remaining on the equilibrium path falls below that of starting afresh in the average community. S and J work if trusted (see also Figure 3).

Let $\pi(l_i, \boldsymbol{\ell}_{-i})$ denote the number of tasks carried out by agent i in state $\boldsymbol{\ell}$ according to the candidate strategies which define the equilibrium. Let the subscripts S,J and F denote the indexes of the respective types. Note that F is assigned no tasks even if he stays given the current realization ($\pi(l_F, \boldsymbol{\ell}_{-F}) = 0$). S does one or two tasks depending on whether F stays or leaves for a new community given the current realization- i.e. $\pi(l_S, \boldsymbol{\ell}_{-S}) \in \{2, 1\}$. J acts as an agent to S: $\pi(l_J, \boldsymbol{\ell}_{-J}) = 1$.

The fact that some individuals may leave on path requires further adaptations of the analysis relative to the baseline HE setting. A first issue is that the transition density from l_i to $\tilde{l}_i$, denoted $\mathcal{G}$, is state dependent.[2] Specifically, suppose v_0 denotes the expected payoff of starting afresh in a new community[3] and $v(l_i, \boldsymbol{\ell}_{-i})$ the value of returning of individual i. The transition density is equal to:

$$\mathcal{G}(\tilde{l}_i, l_i, \boldsymbol{\ell}_{-i}; v, v_0) = \begin{cases} (1-\delta)g(\tilde{l}_i, 0) + \delta g(\tilde{l}_i, l_i) & v(l_i, \boldsymbol{\ell}_{-i}) \geq v_0 \\ g(\tilde{l}_i, 0) & v(l_i, \boldsymbol{\ell}_{-i}) < v_0 \end{cases}. \tag{B1}$$

v_0 in turn depends on the stationary density of the state variable, denoted $\mathcal{F}$ which, being induced by $\mathcal{G}$, is itself endogenous. $\mathcal{F}$ is a solution to the following functional equation

$$\mathcal{F}(\boldsymbol{\ell}; v, v_0) = \int_{[0,1]^3} \mathcal{G}(\tilde{l}_1, l_1, \boldsymbol{\ell}_{-1}, v, v_0)\mathcal{G}(\tilde{l}_2, l_2, \boldsymbol{\ell}_{-2}; v, v_0)\mathcal{G}(\tilde{l}_3, l_3, \boldsymbol{\ell}_{-3}; v, v_0)\mathcal{F}(\tilde{\boldsymbol{\ell}}; v, v_0)\, d\tilde{\boldsymbol{\ell}}. \tag{B2}$$

In addition, the extent of the punishment in the period after a deviation, say $t+1$,

[2]Differently from the main text this object describes the transition density of the i-th component of the state vector. This should not be confused with the related but different object describing the individual's transition density from seniority l_i to $\tilde{l}_i$. These two objects coincide only if individuals never leave or die.

[3]We implicitly assume that the public record of community members can be accessed only after joining so the expectation is taken over the unknown state of the new community.

depends on relative seniority. The most tempted to cheat and return are senior individuals matched with two fully trustworthy agents. Let π^C denote the number of tasks that such senior individual carries over in period $t+1$ in the subgame after cheating in t. Given the above description of the candidate strategies, the value of π^C depends on the realization of the state as well as on the realization of the survival process. So to simplify the notation, in what follows we let $\pi^e(l_S, \boldsymbol{\ell}_{-S})$ denote the expected number of tasks a senior cheater expects to carry out in $t+1$ conditional on returning, which is what matters for incentives, as a function of the vector of realized seniorities in t. The expectation operator accounts for the possibility that the cheated principal(s) do not return in $t+2$ and for the possibility that the parties involved transition to different ranks.[4] We provide a full description of π^C in Appendix B.2.

We define a *Hierarchical Equilibrium with Trading Selectivity and full scale trust* as a set of strategies such that there exists a tuple $(b, v, \pi, \pi^C, \mathcal{G}, \mathcal{F})$ that satisfies $(B1)$, $(B2)$ and the following conditions:

$$v(l_i, \boldsymbol{\ell}_{-i}) = \mathbb{E}_{\tilde{\boldsymbol{\ell}}|\boldsymbol{\ell}} \left[\pi(\tilde{l}_i, \tilde{\boldsymbol{\ell}}_{-i}) w + \delta \max\{v(\tilde{l}_i, \tilde{\boldsymbol{\ell}}_{-i}), v_0\} \right] \tag{B3}$$

$$v_0 = \int_{[0,1]^2} v(0, x, y) \left(\mathcal{F}(0, x, y; v, v_0) / \int_{[0,1]^2} \mathcal{F}(0, \tilde{x}, \tilde{y}; v, v_0)\, d(\tilde{x}, \tilde{y}) \right) d(x, y) \tag{B4}$$

$$\pi(l_S, \boldsymbol{\ell}_{-S}) = 2 \text{ if and only if } v(l_F, \boldsymbol{\ell}_{-F}) \geq v_0 \tag{B5}$$

$$\pi(l_i, \boldsymbol{\ell}_{-i})(s - w) \leq \delta(v(l_i, \boldsymbol{\ell}_{-i}) - v_0) \tag{IC_L}$$

$$\pi(l_i, \boldsymbol{\ell}_{-i})(s - w) \leq \delta \left(v(l_i, \boldsymbol{\ell}_{-i}) - \pi^e(l_i, \boldsymbol{\ell}_{-i}) w - \delta\, \mathbb{E}_{\tilde{\boldsymbol{\ell}}|\boldsymbol{\ell}}[v(\tilde{l}_i, \tilde{\boldsymbol{\ell}}_{-i})] \right) \tag{IC_R}$$

(B3) specifies that the value of each agent depends on the expectation over the next state $\tilde{\boldsymbol{\ell}}$. The operator $\mathbb{E}_{\tilde{\boldsymbol{\ell}}|\boldsymbol{\ell}}$ averages over the convolution of the processes that determine the future states: the survival process and the advancement process. (B4) defines v_0: the expected value of starting afresh in a new community whose state is unknown prior to joining.[5] (B5) requires that the agent with the most seniority in the three person hierarchy receives two trades whenever the newcomer (F) prefers to stay. (IC_L) and (IC_R) represent the no 'cheat and leave' and no 'cheat and return' requirements.

[4] Again recall that the candidate strategies allow F leave on path. So when computing the period payoff in the punishment period $t+1$ we need take into account that in some states F is not trustworthy despite the cooperation phase restarting in $t+2$.

[5] Again note, that unlike in an entry barrier model, starting afresh does not imply a period of 'purgatory' with no trust, but could, in principle, involve being trusted.

The key robustness properties are the following two. First, hierarchical communities are more attractive than institution-free community with full cooperation. This occurs if:

$$v_0 \geq \frac{w}{1-\delta}. \tag{IF}$$

Second, cooperation breaks down in institution-free communities:

$$s + \delta v_0 \geq w + \delta \frac{w}{1-\delta} \tag{B6}$$

Since $s > w$, this condition is implied by (IF). Using numerical methods, we show that the hierarchical equilibrium exists and, in addition, can satisfy the robustness conditions (IF) and (B6).

B.2 Numerical method

In this section we provide an algorithm we used in Matlab to numerically solve for an equilibrium. Computationally, it helps to break the optimization problem into the two cases where the newcomer (F) stays, and where (s)he leaves. Let $L_F \subset [0,1]^3$ be the set of hierarchies where the first player is F and leaves to start anew in alternate community (with expected value v_0) (i.e. $L_F := \{\boldsymbol{\ell} \in [0,1]^3 \colon l_1 = l_F, v(l_1, \boldsymbol{\ell}_{-1}) < v_0\}$.)

The allocation of trade in the punishment one-stage phase, given realization l is given by

$$\pi^C = \begin{cases} 0, & l_1 = l_F \\ 1, & l_1 = l_J, \text{ and } S \text{ is a newcomer} \\ 0, & l_1 = l_J, \text{ and } S \text{ not a newcomer} \\ 2, & l_1 = l_S, \ J \text{ and } F \text{ are both newcomers, and } l_F \notin L_F \\ 1, & l_1 = l_S, \ J \text{ and } F \text{ are both newcomers, and } l_F \in L_F \\ 0, & l_1 = l_S, \ J \text{ is not a newcomer, } F \text{ is a newcomer or } l_F \in L_F \\ 1, & l_1 = l_S, \ J \text{ is newcomer and} F \text{ is not} \\ 0, & l_1 = l_S, \ J \text{ and } F \text{ are not newcomers} \end{cases}$$

We used the following method to solve for HE-TS: first, we solve for v in (B3) forcing the senior (S) and junior (J) members to work; second, we check that S and J are indeed incentivized to stay and work. The key part of the Matlab code solves the dynamic problem (B3), (B4), (B5) by discretizing the state space. Each dimension $[0,1]$ gets `nd` $= 15$ knots, so there are 3375 states in the state space. On the grid of (b, δ) ($b \in [.1, .9]$ with step .1, $\delta \in [.7, .98]$ with step .02), we do the following:

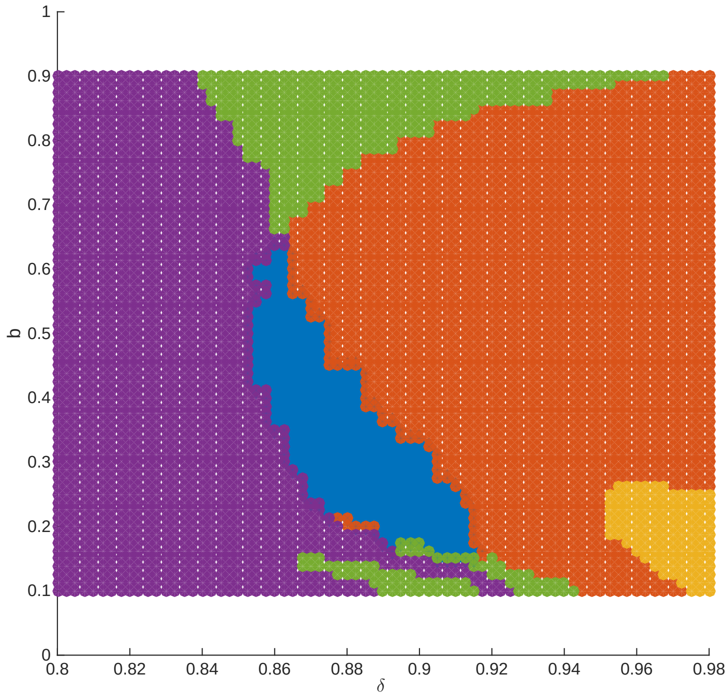
1. Construct state transition probability matrix `P` from g on the grid with `nd` $= 15$. For each action, `P` is 3375×3375.
2. Construct the initial stage reward matrix `R`: it assigns stage payoff $2w$ to S and stage payoff w to J, and stage payoff zero to F.[6]
3. Solve the dynamic optimization problem (B3) by iterating the value function.
4. Update `R` to take into account that S's stage payoff is w in a state where F leaves. Update `P` to take into account that F leaves in some states of the world. Go to step 3 if new `R` and `P` are different from the old `R` and `P`; otherwise, go to step 5.
5. Check whether agents in an institution-free community prefer to cheat defect to the hierarchy $(v_0 \geq \frac{w}{1-\delta})$ (note that this turns out not to be restrictive.)
6. Check if IC conditions (IC_L) and (IC_R) hold for any s. If any of the conditions fail, record the non-existence of HE-TS.

B.3 Existence

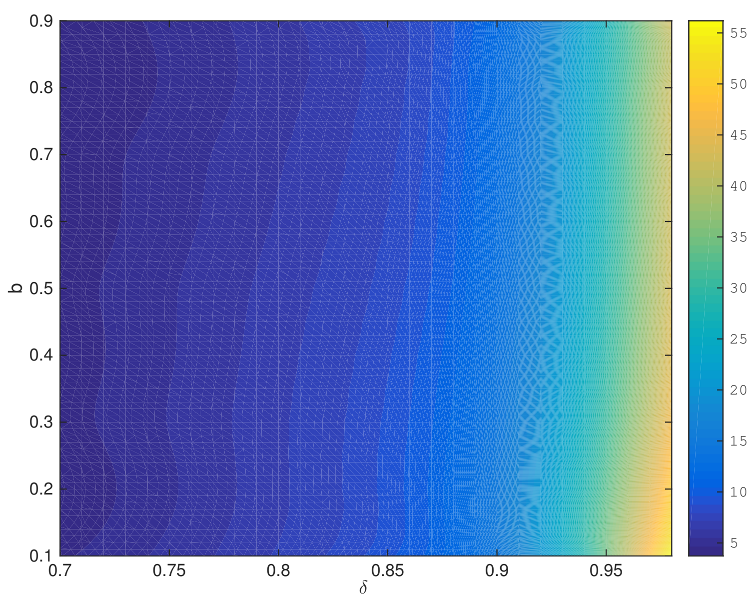
Figure 4a shows the existence region of the hierarchical equilibrium, depending on the level of patience δ and the extent of social mobility b. Notice that there are two main regions, corresponding to the two main incentive constraints (IC_L and IC_R). First, (IC_L) is satisfied for high degrees of a patience, as one might expect, but also at intermediate levels of social mobility (the red and blue regions). Recall that the binding constraint to not cheat and leave is on higher status agents being trusted by multiple partners of lower status. The presence of some social mobility that allows these partners to be promoted in the future makes them less desirable marks for a higher status cheater to exploit.

In contrast to (IC_L), (IC_R) is satisfied for lower levels of patience (blue and purple). This is because if the agent is patient, he can more easily tolerate the one-period punishment. The resulting existing region is the blue area sustained by intermediate levels of both patience δ and social mobility b. Notice that violations of the condition IF (in orange) are nested inside the same parameter space that also violates IC_R– so there is no additional restrictions on patience or social mobility to obtain institutional robustness.

[6] Note that here we force S and J to work on the equilibrium path.



(a) Existence region. Horizontal axis is patience (δ), vertical axis is social mobility (b), $w = 1$. Blue diamond: equilibrium exists for some $s > w$. Red: (IC_R) is violated; purple: (IC_L) is violated; green: (IC_L) and (IC_R) are violated; orange: (IC_R) and (IF) violated.



(b) Payoff of an average member of a hierarchical community with trading selectivity in (δ, b) space.

Figure 4

Thus, hierarchies can not only be robust to other institutions, but their presence can also make it impossible to sustain alternative institutional arrangements, when agents have the option to cheat and leave.

Finally, Figure 4b shows how the value for an average member of the hierarchy changes with social mobility and patience.[7] The long-term payoff is higher for longer lived members with lower social mobility (low b).

[7] That is, the average value $v(l)$ where the expectation is taken over the stationary distribution of l's.

C Appendix: Additional Applications

The Birth of Hierarchy in Early Human Societies Prior to the introduction of settled agriculture around 12,000 years ago, human societies consisted of small groups and population densities as a whole were relatively low (Boix and Rosenbluth, 2014, Keeley, 1988). They also tended to be egalitarian (Boix and Rosenbluth (2014)). Inequality increased, and average health appeared to worsen, however, after the introduction of agriculture (Boix and Rosenbluth, 2014, Bowles and Choi, 2013), a period that also saw an increase in population densities and the emergence of social hierarchies (eg Boehm, 1999, Price and Brown, 1985).[8] Some scholars argue that egalitarianism was pinned down by the "countervailing power" of weaker individuals– particularly their ease of exit, should any 'self-aggrandizing' individual seek to dominate (Woodburn, 1982, Knauft, 1991, Boehm, 1999, Nowak and Sigmund, 2005, Sterelny, 2013, Seabright, 2013, Bowles and Choi, 2013).

An important body of work links the emergence of hierarchies and within-group inequality to the development of settled agriculture. Many existing explanations relate to heterogeneity among individuals. For example, Dow and Reed (2013) present a theoretical model where geographical heterogeneity in agricultural suitability generates inequality, as early movers to desirable plots of land collude to exclude latecomers from property ownership.[9] In contrast, our model suggests that such heterogeneity is not necessary for inequality to develop. Any exogenous change that makes it more attractive to cheat and leave, *even if this affects all individuals equally*, could lead to the development of hierarchies. For example, hierarchies might simply emerge from technological improvements that reduce travel costs, or increases in population density, that make alternative trading partners and communities more accessible and thus make it easier to cheat and leave any particular individual or group. Going back to the model, if we think of the membership fee in the identity investment equilibria m as being determined by the distance needed to travel between them, having cheated a partner, increased proximity then between markets (an m lowered below $\underline{m}$) can lead to a breakdown of cooperation in those internally-egalitarian communities. In contrast, as we have seen, cooperation

[8] Summarizing a large body of anthropological and archaeological work, the anthropologist Robert Boehm (1999)[pg3-4] states: "before twelve thousand years ago, humans basically were egalitarian. They lived in what might be called societies of equals, with minimal political centralization and no social classes. Everyone participated in group decisions and outside the family there were no dominators."

[9] See also Boix (2015), who argues that agricultural technology favoured some agents more than others, leading the latter to become bandits and then specialists in violence. Mayshar, Moav, Neeman, and Pascali (2015) point to the extent to which agricultural staples in some places required storage and thus were appropriable in leading to the potential for hierarchies and state formation.

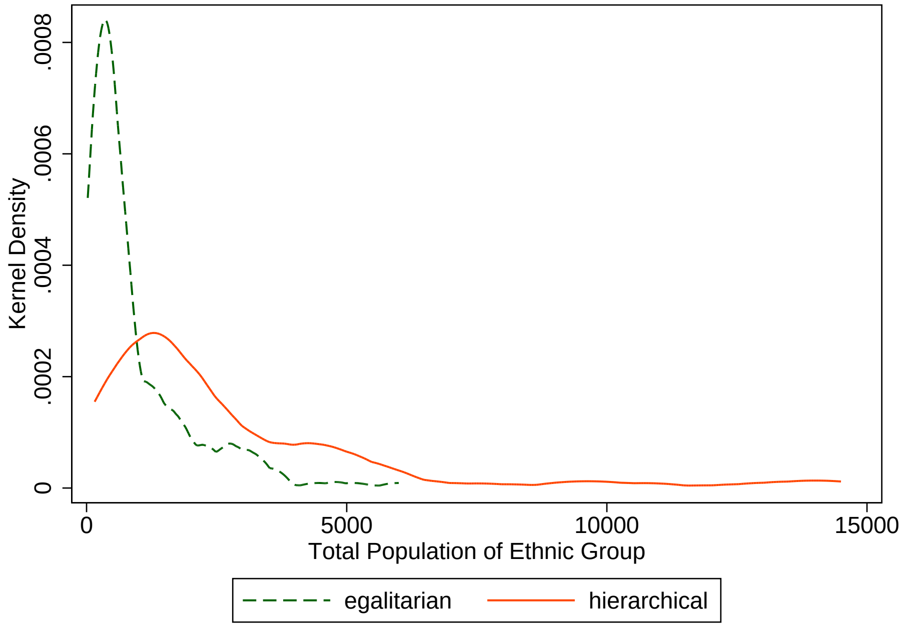


Figure 5: **Egalitarian Hunter-Gatherer Ethnic Groups are Smaller**
Kernel densities of ethnic group size are based upon the Binford (2001) dataset of 339 hunter-gatherer societies described at first contact with anthropologists. This viewed as a comprehensive and expanded version of those mentioned in Murdock's *Ethnographic Atlas*. *Egalitarians* include all societies classified as "generic hunter-gatherers" or "generic hunter-gatherers with institutionalized leaders". *Hierarchical* are all societies classified as "wealth-differentiated hunter-gatherers" or "stratified or characterized by elite and privileged leaders".

in hierarchical communities could survive the reduction, or even the elimination of the costs of inter-community travel.

Figures 5 and 6 reveal a positive relationship between hierarchy and population density in a "comprehensive" dataset of modern (non-agricultural) hunter-gatherer societies, as observed upon their first contact with Europeans (Binford, 2001)[pg.117]. Consistent with our theory, egalitarian hunter-gatherer ethnic groups tend to be much smaller, even if we include within them those with some form of institutionalized leadership (Figure 5). Further, (Figure 6) while egalitarian institutions survive with low population densities, even among hunter-gatherer societies, ethnic groups with wealth distinctions and social stratification begin to be more prevalent when population densities exceed as little as 1 person per square kilometer.

Our model also sheds new light on a lingering paradox: that hierarchies and inequalities *precede* the development of early states with coercive power that might justify having

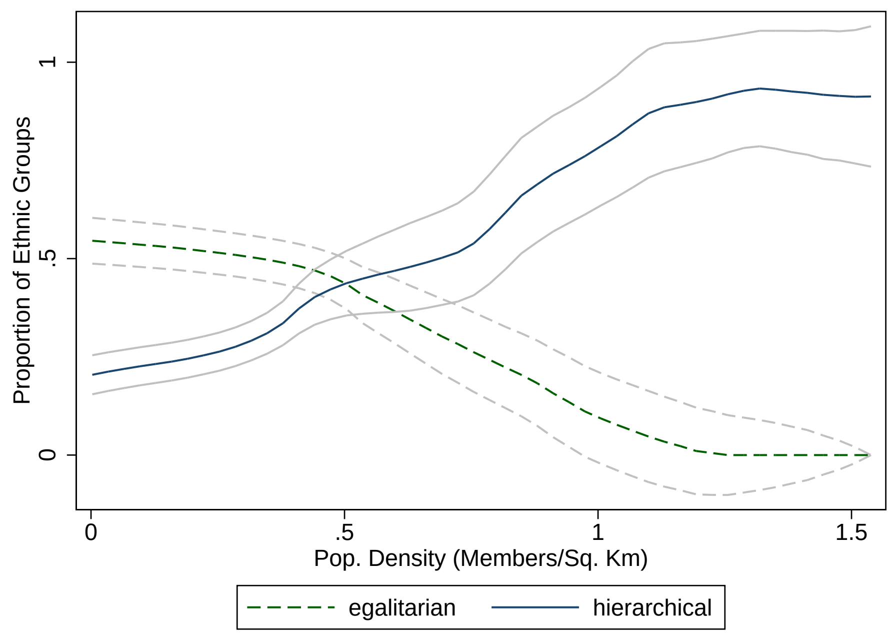


Figure 6: **Egalitarian Groups Do Not Survive as Population Densities Rise**
Local moving averages (epanechnikov kernel, ROT bandwidth) showing the proportion of societies that are egalitarian or hierarchical at different levels of ethnic group population densities. These are based upon the Binford (2001) dataset of 339 hunter-gatherer societies described at first contact with anthropologists. This viewed as a comprehensive and expanded version of those mentioned in Murdock's *Ethnographic Atlas*. *Egalitarians* include all societies classified as "generic hunter-gatherers" or "generic hunter-gatherers with institutionalized leaders". *Hierarchical* are all societies classified as "wealth-differentiated hunter-gatherers" or "stratified or characterized by elite and privileged leaders".

large groups for mutual defence, or indeed might result in the security of property rights that would allow 'self-aggrandizing' individuals to assert economic distinction in the first place (Seabright, 2013).[10] The hierarchical equilibria that emerge in our model address the question of why elites emerge: while egalitarian societies are favoured when distances or travel costs are great and when the tasks individuals can perform for one another are also extremely limited, hierarchies are favoured when these capacity constraints are relaxed. Furthermore, in environments where trade can be concentrated with senior members, hierarchies can be designed that are better able to sustain cooperation in large groups, and are robust to *countervailing power*- coalitional deviations. The development of social distinction also appears natural and intuitive- the hierarchies that emerge

[10]In his survey of archaic states, Norman Yoffee (2005)[pg.35], based upon archaeological evidence such as differing sizes of residences, the distribution of artifacts, and mortuary furnishings, concludes that"no prehistoric trajectory to any state fails to contain indications of significant economic inequality or the potential of such inequality well before the appearance of anything that might be called a state."

transfer surplus from newcomers to incumbent (returning) members of society, yet sustain incentives for newcomers to themselves return. Our model also explains the key question of institutional selection: hierarchies can emerge spontaneously that dominate and lead to the breakdown in other institutional arrangements, and even undermine groups that had previously been able to maintain such cooperation without institutions. Thus, even without any additional assumptions that hierarchies enjoy a technological advantage in coercive organization, hierarchies may emerge spontaneously that sustain cooperation, and may have provided the seed of early political states.[11]

Citizenship in the Roman Republic, Carthage and Hellenistic Greece Another novel aspect of our model is that, unlike social networks which are often person-specific, social hierarchies are *impersonal* and the patterns of cooperation in a social hierarchy can survive beyond the individuals that populate them. In endogenous social network theory (eg Jackson, 2003, Bloch, Genicot, and Ray, 2008)), ties are based upon an individual's trading history. Instead, in our framework, centrality in the network derives from an individual's office or status. Agents *salute the rank* and can trust senior individuals even if they have never met before. Thus, the social structure can out-live an individual agent or hereditary lineage. Indeed, the development of institutions that support such impersonal relations has been seen as critical for the historic expansion of trade (Greif, 2000).

Furthermore, the impersonal hierarchy in our model is akin to a concept of citizenship which was also novel in the classical world, that of the Roman Republic (ca. 2nd-3rd centuries BC).[12] Rome's contemporaries, like Carthage and the Greek city-states, were organized along traditional ethnic lines that mimic the identity investment equilibrium. Citizenship derived mainly from ethnic lineage. Citizenship in Athens, for example, originally derived from membership in one of four tribes (Aristotle, 335BC)[pg.56,65]. Citizens had specific privileges, many that built upon trust and specialization, that included relatively equitable economic opportunities and political rights, including in a number of prominent cases like Athens, democracy (Ober, 2015)[chp 1]. However, there was systematic exclusion from these rights and opportunities of all outside the specific

[11] The question of how property rights, inequalities and ultimately coercive elites emerged and were sustained despite strong egalitarian norms and the possibility for coalitions to seize and redistribute any unequal capital accumulation has long been a puzzle (Sterelny, 2013). Among the other ways that this issue has been addressed has been the assumption of a behavioural 'lag' in cooperative social norms– that individuals cooperate with the nascent elites and allow them to take advantage of this (Seabright, 2010, Richerson and Boyd, 2001, Sterelny, 2013).

[12] We are grateful to Ian Morris for pointing out these parallels with our model.

ethnic group of the city, with 'foreigners' including citizens of other city-states, ex-slaves and, in the case of colonies, even the original inhabitants.

Incentives in the towns of the Roman Republic, in contrast, arguably paralleled the hierarchical equilibrium. They were remarkable for the time by allowing ethnic groups with no genetic relations to Rome, and living far away from the mother city to become *Roman citizens*. There were a number of ranks: slaves, citizens, junior and senior office-holders, culminating in senators, each enjoying higher levels of trust and privileges. Remarkably for the time, there was also gradual but probabilistic advancement at each rank. Many slaves had a good chance of gaining freedom and full Roman citizenship in their lifetime, so much so that by the second century CE, the majority of citizens are believed to have had slave ancestry (Beard, 2015)[pg.67]. This mobility extended to the top: the elite senatorial class itself became a multi-ethnic body (Beard, 2015)[pg.68].

This impersonal institutional structure did not limit Rome by the size of any particular ethnic group and instead allowed Rome to accommodate a large, geographically dispersed and rapidly expanding population. In fact, the comparison between the "avaricious" Greeks and the "generous" Romans in their bestowal of citizenship has been credited both by contemporary observers and modern historians for Rome's relative ability to scale, and its ultimate dominance over the Hellenistic Mediterranean (Gauthier, 1974, Eckstein, 2008, Beard, 2015).[13]

Our model suggests that the Carthaginians' and Greeks' unwillingness to admit members that could have been trustworthy and their attendant inability to scale relative to the Romans may have reflected the historic incentive problems they faced in maintaining cooperation given the presence of other proximate city-state communities in the Hellenis-

[13] A famous stelae from 215 BC illuminates a contemporary Greek perspective on Roman institutions. It records an exchange to King Philip V of Macedon and the semi-independent Greek city of Larisa in Thessaly. In response to a previous letter suggesting that the Larisans should overcome their recent war-related depopulation by admitting new citizens from among its inhabitants, the Larisans had first admitted and then revoked the citizenship of more than 200 new inhabitants, the majority from Thessaly itself. In response, the King wrote:

> I hear that those who were granted citizenship in accordance with the letter I sent you and your decree, and whose names were inscribed (on the stele) have been erased. If this has happened, those who have advised you have ignored the interests of your city and my ruling. That it is much the best state of affairs for as many as possible to enjoy citizen rights, the city to be strong and the land not to lie shamefully deserted, as at present, I believe none of you would deny, and one may observe others who grant citizenship in the same way. Among these are the Romans, who when they manumit their slaves admit them to the citizen body and grant them a share in the magistracies, and in this way have not only enlarged their country but have sent out colonies to nearly 70 places." (Austin, 2006)[pg 157-159].

tic Mediterranean. In contrast, the adoption by the Romans of a system of *impersonal* hierarchy with social mobility may have been a contributing factor in the Republic's rapid ascendancy and subsequent dominance.[14]

Online Communities Open Source Software (OSS) communities and internet discussion groups are increasingly important venues for exchange Raymond (2001), Lerner and Tirole (2005), Shah (2006). As early as 2006 there were over one hundred thousand open source projects with over a million registered users. A central feature of online communities at first seems surprising from the perspective of standard incentive theory: individuals regularly contribute to a public good despite the lack of monetary incentives, large group size and the relative anonymity of the interactions. A closer look at these communities reveals that it is not purely public-spiritedness that motivates users however. Many groups have formal or informal hierarchies. In OSS communities, there is often a group of senior members known as "committers" who have the authority to incorporate both their own code into the project, as well as that of others. In addition, there is often an informal hierarchy, where more senior members will help answer questions for other senior members, but not junior members. In other internet discussion groups, it is common to show the date a user joined beside their posts.[15] Other types of designations (such as top reviewers for Amazon.com, or Most Valuable Professionals on Microsoft Forums) also help distinguish the senior members of a community. Our analysis helps to understand these phenomena.

Consider how OSS projects fit into the model. Even though in principle individual contributions may be publicly observable in many online communities, in practice it requires a substantial investment to learn about the quality of their work. One might need to carefully read a programmer's code. Some aspects of support and collaboration in OSS projects take place in private correspondence. It may thus be difficult for outsiders to ascertain the quality of interactions others are having in the community, and one may only learn about an individual's behavior by trying to adopt the code or otherwise having a close interaction with someone.

Our model applies to a setting where programmers select whether to become involved

[14]A traditional historical view credits Rome's military-oriented culture, including the glorification of warfare. As Eckstein (2008) points out, however such cultures were pervasive among contemporary states, particularly amongst the successor-states of Alexander's empire, and military advances were rapidly shared and adopted across communities.

[15]As Microsoft's Office Forum web page indicates, "the information about your activity in the Community (such as how many posts you have contributed...) gives others a sense of your trustworthiness and a way to gauge how valuable your comments might be."

in a particular project. Programmers may write code for different purposes. Individuals associated with the project have needs arise which may be met using code that others have written, and they may choose whether or not to invest in reading and trying to use the code (as opposed to writing from scratch or finding other sources). They may then have support questions. The author of the code can then choose whether to support the user, answer questions, etc.[16] The "scale" of trust on the part of the user can be interpreted in several ways: for example, the user might make use of only the basic features of the code, or the user might invest rewriting portions himself. The user observes whether the ex ante quality was high, and also observes whether support was provided. The author gets higher utility the more the user relies on the author's original code (rather than rewriting it), because competing versions of the same code are thereby avoided. In such settings, the observed seniority-based hierarchies of trust and support can be understood as institutions that engender cooperation, and perhaps the puzzle of how OSS and online discussion groups incentivize trust is less puzzling in light of the finding that hiearchies can be powerful forces in favor of cooperation.

[16]Some of the author's choices may be made *ex ante*, such as deciding on the extent of documentation and code quality, modularity/adaptability, etc.

References

Akerlof, G. A., and R. E. Kranton (2000): "Economics and identity," *Quarterly Journal of Economics*, 115(3), 715–53.

Aristotle (335BC): *Politics*. Chicago U. Press, Chicago, translated by Carnes Lord, 2013 edn.

Arrow, K. J. (1974): *The limits of organization*. Norton, New York, NY, 1st edn.

Austin, M. (ed.) (2006): *The Hellenistic World from Alexander to the Roman Conquest. A selection of ancient sources in translation.* Cambridge University Press, 2nd edn.

Barkow, J. H. (1973): "Muslims and Maguzawa in North Central State, Nigeria: an Ethnographic Comparison," *Canadian Journal of African Studies*, 7(1), 59–76.

——— (1974): "Evaluation of Character and Social Control among the Hausa," *Ethos*, 2(1), 1–14.

Bates, R., A. Greif, and S. Singh (2002): "Organizing violence," *Journal of conflict resolution*, 46(5), 599–628.

Beard, M. (2015): *SPQR: A History of Ancient Rome*. Liveright.

Besley, T., and T. Persson (2009): "The origins of state capacity: property rights, taxation and politics," *American Economic Review*, 99(4), 1218–1244.

Binford, L. R. (2001): *Constructing Frames of Reference: An Analytical Method for Archaeological Theory Building Using Hunter-Gatherer and Environmental Data Sets.* University of California Press, Berkeley.

Bisin, A., and T. Verdier (2011): "The Economics of Cultural Transmission and Socialization," in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. O. Jackson, chap. 9. North Holland.

Bloch, F., G. Genicot, and D. Ray (2008): "Informal insurance in social networks," *Journal of Economic Theory*, 143, 36–58.

Board, S. (2008): "Relational contracts and the value of loyalty," mimeo, UCLA.

Boehm, R. (1999): *Hierarchy in the Forest*. Harvard University Press.

Boix, C. (2015): *Political Order and Inequality*. Cambridge University Press.

Boix, C., and F. Rosenbluth (2014): "Bones of Contention: The Political Economy of Height Inequality," *American Political Science Review*, 108(1), 1–22.

Bowles, S., and J.-K. Choi (2013): "Coevolution of farming and private property during the early Holocene," *Proceedings of the National Academy of Sciences*, 110(22), 8830–3335.

Boyd, R., and P. J. Richerson (1994): "The evolution of norms: an anthropological view," *Journal of Theoretical and Institutional Economics*, 150(1), 72–87.

Carmichael, H. L., and W. Macleod (1997): "Gift giving and the evolution of cooperation," *International Economic Review*, 38, 485–508.

Cohen, A. (1969): *Customs and politics in Urban Africa: a study of Hausa migrants in Yoruba towns*. University of California Press, Berkeley.

Dixit, A. K. (2003): "Trade expansion and contract enforcement," *Journal of Political Economy*, 111(6), 1293–1317.

Doepke, M., and F. Zilibotti (2013): "Culture, Entrepreneurship and Growth," working paper 19141, NBER.

Dow, G. K., and C. G. Reed (2013): "The Origins of Inequality: Insiders, Outsiders, Elites and Commoners," *Journal of Political Economy*, 121(3), 609–641.

Eckstein, A. M. (2008): *Rome enters the Greek East : from anarchy to hierarchy in the Hellenistic Mediterranean, 230-170 BC*. Blackwell Pub., Malden, MA.

Friedman, E., and P. Resnick (2001): "The social cost of cheap pseudonyms," *Journal of Economics and Management Strategy*, 10(2), 173–199.

Fryer, R. G. (2002): "An economic approach to cultural capital," mimeo, University of Chicago.

Gauthier, P. (1974): "Générosité romaine et 'avarice' grecque: sur l'octroi du droit du cité," in *Melanges d'histoire ancienne offerts á William Seston*, pp. 207–215. de Boccard, Paris.

Genicot, G., and D. Ray (2003): "Endogenous Group Formation in Risk-Sharing Arrangements," *Review of Economic Studies*, 70(1), 87–113.

Ghosh, P., and D. Ray (1996): "Cooperation in community interaction without information flows," *Review of Economic Studies*, 63(3), 491–519.

Gibbons, R. (2005): "What is economic sociology and should any economists care?," *Journal of economic perspectives*, 19(1), 3–7.

Greenberg, J. (1947): "Islam and clan organization among the Hausa," *Southwestern Journal of Anthropology*, 3(3), 193–211.

Greif, A. (1993): "Contract enforceability and economic institutions in early trade: the Maghribi traders' coalition," *American Economic Review*, 83(3), 525–547.

——— (2000): "Impersonal exchange and the origin of markets: from the community responsibility system to individual legal responsibility in pre-modern Europe," in *Communities and markets*, ed. by M. Aoki, and Y. Hayami, chap. 2. Oxford University Press, Oxford.

Hannan, M. T., L. Polos, and G. R. Carroll (2007): *Logics of Organization Theory: Audiences, Code, and Ecologies*. Princeton University Press, Princeton.

Hermalin, B. (2010): "Leadership and corporate culture," in *Handbook of Organizational Economics*, ed. by R. Gibbons, and J. Roberts. Princeton University Press.

Iannaccone, L. R. (1992): "Sacrifice and stigma: reducing free-riding in cults, communes and other collectives," *Journal of Political Economy*, 100(2), 271–291.

Jackson, M. O. (2003): "A survey of models of network formation: stability and efficiency," in *Group formation in economics: networks, clubs and coalitions*, ed. by G. Demange, and M. Wooders, Cambridge, UK. Cambridge University Press.

Kandori, M. (1992): "Social norms and community enforcement," *Review of Economic Studies*, 59, 63–80.

Keeley, L. H. (1988): "Hunter-Gatherer Economic Complexity and 'Population Pressure': A Cross-Cultural Analysis," *Journal of Anthropological Archaeology*, 7(4), 373–411.

King, L. D. (2006): *Africa and the Nation State: State formation and identity in ancient Egypt, Hausaland and Southern Africa*. The Edwin Mellen Press, Lewiston.

Klein, B., and K. B. Leffler (1981): "The role for market forces in assuring contractual performance," *Journal of Political Economy*, 89(4), 615–41.

Knauft, B. (1991): "Violence and sociality in human evolution [and comments and replies]," *Current Anthropology*, 32(4), 391–428.

Kranton, R. E. (1996a): "The formation of cooperative relationships," *Journal of Economics, Law and Organisation*, 12(1), 214–33.

——— (1996b): "Reciprocal exchange: a self-sustaining system," *American Economic Review*, 86(4), 830–851.

Kreps, D. M. (1990): "Corporate culture and economic theory," in *Perspectives on Positive Political Economy*, ed. by J. E. Alt, and K. A. Shepsle, pp. 90–143. Cambirdge University Press, Cambridge.

Last, M. (1979): "Some economic aspects of conversion in Hausaland," in *Conversion to Islam*, ed. by N. Levitzion, chap. 13, pp. 236–246. Holmes and Meier, New York.

——— (1999): "History as religion: de-constructing the Magians 'Maguzawa' of Nigerian Hausaland," in *L'invention religieuse en Afrique*, ed. by J.-P. Chretien, C.-H. Perrot, G. Prunier, and F. Raison-Jourde, chap. 11. Karthala.

Lerner, J., and J. Tirole (2005): "The economics of technology sharing: open source and beyond," *Journal of Economic Perspectives*, 19(2), 99–120.

Lindsey, J., B. Polak, and R. Zeckhauser (2003): "Free love, fragile fidelity and forgiveness: rival social conventions under hidden information," mimeo.

Mayshar, J., O. Moav, Z. Neeman, and L. Pascali (2015): "Cereals, appropriability and hierarchy," mimeo, Warwick.

McMillan, J. (2002): *Reinventing the Bazaar: a natural history of markets*. W.W. Norton & Co., New York.

North, D., J. Wallis, and B. Weingast (2009): *Violence and social orders: a conceptual framework for interpreting recorded human history*. Cambridge University Press.

Nowak, M. A., and K. Sigmund (2005): "Evolution of indirect reciprocity," *Nature*, 437, 1291–1298.

Ober, J. (2015): *The Rise and Fall of Classical Greece*. Princeton University Press.

Price, T. D., and J. A. Brown (1985): *Prehistoric Hunter-Gatherers: The Emergence of Cultural Complexity*. Academic Press, Orlando.

Ramey, G., and J. Watson (2001): "Bilateral trade and opportunism in a matching market," *Contributions to theoretical economics*, 1(1).

Raymond, E. S. (2001): *The Cathedral and Bazaar: musings on linux and open source by an accidental revolutionary*. O' Reilly, Cambridge, MA.

Richerson, P., and R. Boyd (2001): *The evolution of subjective commitment to groups: A tribal instincts hypothesis*pp. 186–220. Russell Sage, New York, NY.

Salamone, F. A. (2010): *The Hausa of Nigeria*. The University Press of America.

Seabright, P. (2010): *The company of strangers: a natural history of economic life*. Princeton University Press, Princeton.

——— (2013): *The Birth of Hierarchy*.chap. 5, pp. 109–117, in Sterelny, Joyce, Calcott, and Fraser (2013).

Shah, S. (2006): "Motivation, governance and the viability of hybrid forms in open source software development," *Management Science*, 52(7), 1000–1014.

Shapiro, C., and J. E. Stiglitz (1984): "Equilibrium unemployment as a worker discipline device," *American Economic Review*, 74(3), 433–444.

Sobel, J. (1985): "A theory of credibility," *Review of Economic Studies*, 52(4), 557–73.

——— (2006): "For better or forever: formal versus informal enforcement," *Journal of Labor Economics*.

Sterelny, K. (2013): *Life in Interesting Times: Cooperation and Collective Action in the Holocene*, in Sterelny, Joyce, Calcott, and Fraser (2013).

Sterelny, K., R. Joyce, B. Calcott, and B. Fraser (eds.) (2013): *Cooperation and its evolution*MIT Press, Cambridge, MA.

Van den Steen, E. (2005): "On the Origin of Shared Beliefs (and Corporate Culture)," MIT Sloan working paper 4553-05.

Watson, J. (1999): "Starting small and renegotiation," *Journal of Economic Theory*, 85, 52–90.

White, H. C. (1970): *Chains of opportunity: system models of mobility in organizations.* Harvard University Press, Cambridge, MA.

Woodburn, J. (1982): "Egalitarian Societies," *Man*, 17(3), 431–451.

Woodruff, C. (1998): "Contract enforcement and trade liberalization in Mexico's footwear industry," *World Development*, 26(6), 979–991.

Yoffee, N. (2005): *Myths of the Archaic State: Evolution of the Earliest Cities, States and Civilizations.* Cambridge University Press, Cambridge.