Endogenizing fractionalization*

PETER T. LEESON
Department of Economics, George Mason University

Abstract: In this paper I identify fractionalization as a consequence rather than solely a cause of poor institutions. I investigate how heterogeneous agents in precolonial Africa relied on social distance-reducing signals to make trade with one another possible. I then show how colonial institutions created noise in these signals, inhibiting widespread cooperation. By stifling trade between diverse agents, colonial institutions contributed to Africa’s poor economic growth.

1. Introduction

Why are some nations rich and other nations poor? While economists are far from having answered this question definitively, two forces in particular seem to play a critical role in determining economic performance: institutions and fractionalization.

The recent literature examining this question is divided in its emphasis on these forces. One approach emphasizes the institutional differences that distinguish different countries. In their important work Acemoglu et al. (2002, 2001) divide European colonization into two groups: colonization that occurred in places like America, Canada, Australia, New Zealand, Hong Kong and Singapore, and colonization that took place in regions like Africa.

According to these authors, colonization in the former group was characterized by the future-looking European settlement that established institutions conducive to long-term growth. In contrast, the latter group was part of the ‘scramble for Africa’ in which colonial powers sought to get in, extract as much possible from the indigenous population, and get out. The establishment of poor institutions in this case led to long-term economic problems in African countries.

The other dominant approach to economic growth disparities emphasizes the importance of fractionalization in preventing economic progress. According to this literature the presence of numerous, socially distant individuals creates poor-quality governments in nations where it is prominent (see for instance, Alesina and Spolaore, 2003; Easterly and Levine, 1997; Alesina et al., 1999; Alesina et al., 2003; La Porta et al., 1999). For instance, high fractionalization in

* I am grateful to Peter Boettke, Tyler Cowen, Chris Coyne, Melissa Thomas, Richard Wagner and three anonymous reviewers for helpful comments and suggestions. The financial assistance of the Mercatus Center and the Olofsson Weaver Fellowship is also acknowledged.
sub-Saharan Africa prevents the state from performing necessary functions like providing public goods, and generates ethnic conflict.

The majority of struggling nations exhibit both poor institutions and high fractionalization. Recognition of this phenomenon begs the question, how are institutions and fractionalization connected? The fractionalization approach attempts to answer this question by explaining poor institutions as the outcome of the presence of socially heterogeneous individuals. However, little work has addressed the connection between institutions and fractionalization from the institutional approach. My aim here is to do this.

Recent work by Easterly (2001) points to the importance of institutions in mitigating the problems typically associated with fractionalization. He shows that sufficiently good institutions eliminate the negative effect that the presence of socially diverse individuals has on economic growth. This insightful work is an important first step in demonstrating that social heterogeneity *per se* is not the ultimate cause of economic strife.

What matters for progress is the ability of individuals to realize the gains from widespread exchange. What should therefore concern us is not the probability that two randomly selected agents who interact will be socially distant (the standard measure of fractionalization), but rather whether or not two socially heterogeneous agents can peacefully exchange. I define fractionalization in these terms rather than as it has traditionally been defined. Non-fractionalized agents are able to reap the benefits from trade despite being socially distant. Fractionalized agents, in contrast, are unable to do so and instead interact predominantly with those they are very close to and thus know very well.

In this spirit, this paper is interested in identifying the mechanism working through bad institutions that prevents agents from realizing the gains from widespread exchange. Building on Easterly (2001) who considers the role of institutions in mitigating fractionalization, I consider the ‘flip-side’ of Easterly’s thesis and examine the importance of institutions in *exacerbating* fractionalization. Specifically, I consider the arrow of causation running from low-quality government (i.e., one with poor institutions) to increased fractionalization, which has been largely neglected in favor of that running from fractionalization to low-quality government. These alternative approaches are depicted below in Figure 1.

Figure 1. Alternative Sequences of Causation

\[
\text{Fractionalization} \rightarrow \text{Bad Institutions} \rightarrow \text{Economic Deterioration} \\
\text{Traditional Approach}
\]

\[
\text{Bad Institutions} \rightarrow \text{Fractionalization} \rightarrow \text{Economic Deterioration} \\
\text{My Approach}
\]
Like those before it, this analysis is concerned specifically with the case of Africa. I do not cover all the nuances or angles of fractionalization or Africa’s growth problem. These issues are too vast and complicated to be addressed in their entirety here. For instance, I do not address ‘ethno-genesis’ – the arbitrary creation of new ethnic groups in Africa, questions of population density, and many other factors that may have directly impacted fractionalization and low growth in Africa.

Similarly, in discussing colonial institutions, I do not discuss the direct effects that colonial institutions may have had on fractionalization or growth. Such direct effects of colonization – for instance the arbitrary delineation of new borders, the creation of new tribes in parts of Africa, and the ‘divide and rule’ methods of many colonizers – are already well-known and lie outside the scope of my analysis. Instead, I aim to offer insight regarding the equally important but harder to identify and thus overlooked indirect effects of colonial institutions on the ability of Africans to peacefully trade.

Unfortunately, sufficient quantitative data regarding interactions between socially heterogeneous agents in pre-colonial Africa is unavailable. This severely hampers empirical assessment, since it is not possible to perform econometric analyses. To circumvent this problem, I rely exclusively on historical accounts, which provide important insights into the pre-colonial and colonial environments of widespread trade.

My analysis admittedly draws on a ‘patchwork’ of historical examples in its approach, which are not applicable to all parts of, or peoples within, Africa. The historical discussions contained here are therefore unavoidably general and necessarily gloss over or ignore many important items. With this cautionary note in mind, it is nevertheless my hope in drawing upon these examples to offer insight regarding the connection between bad institutions, fractionalization, and low growth in Africa.

My thesis is as follows: Poor institutions are as much the cause as they are the effect of fractionalization. ‘Risk of expropriation’ (Acemoglu, Johnson, and Robinson, 2001: 1235) captures only part of what it means for an institution to be harmful. Equally important is the extent to which various types of institutions interfere with the signaling mechanism that makes exchange between heterogeneous agents possible. Pre-colonial agents used signals that reduced the social distance between sender and receiver to convey their credibility to outsiders they wanted to trade with. By conveying their sender’s credibility, these signals enabled widespread exchange. I consider three specific pre-colonial arrangements that functioned in this capacity: the individual’s relationship to authority and community, property usage, and religious practice/associations.

Colonial-created institutions of the Native Authority, property law, and religious policy interfered with the signals that enabled peaceful interaction between socially heterogeneous individuals. Specifically, they severed the communication mechanisms between the diverse inhabitants of Africa. As a
result, socially distant individuals were prevented from realizing the gains from trade with one another and conflict emerged between them. The resulting inability of individuals to capture the widespread gains from trade contributed to Africa’s poor economic performance.

The remainder of this paper is organized as follows: Section 2 provides a theoretical framework for understanding the signaling mechanism that makes peaceful exchange between heterogeneous agents possible. Section 3 applies this framework to Africa by considering three arrangements operating among pre-colonial Africans for this purpose. Section 4 demonstrates how ‘bad’ institutions destroy the signaling mechanism that enables trade between heterogeneous agents and in doing so increase fractionalization. Section 5 shows how the colonial institutions of the Native Authority, land policy, and religious laws incapacitated the pre-colonial arrangements for signaling considered in Section 3. Section 6 concludes.

2. Signaling to expand trade

Where formal enforcement is weak or absent, as it was in much of pre-colonial Africa, individuals must rely upon informal mechanisms for coping with the potential situations of conflict they confront. Historical examples of this abound and a substantial and growing literature continues to bring instances of such mechanisms to our attention (for only a few see, Greif, 1993; Greif et al., 1994; Milgrom et al., 1990; Kranton, 1996; Landa, 1994; Zerbe and Anderson, 2001; Benson, 1989). Most of this literature demonstrates that inside small homogeneous social groups, informal mechanisms of multilateral punishment (for instance, boycott) can support cooperation. However, as the number of individuals to which this mechanism is to apply expands, multilateral punishment breaks down. People simply become too numerous to permit information regarding their histories to be effectively communicated to everyone involved. Communication in this case can become prohibitively costly or outright impossible if the population is large enough.

This problem is exacerbated where the operation of multilateral punishment is to be applied to socially distant individuals. Besides making it more costly to communicate with others, social heterogeneity makes it more difficult for individuals to converge upon necessary social norms that dictate both what constitutes cheating and how cheaters are to be punished. As Greif (2002) points out, the problem of social heterogeneity can be overcome if heterogeneity exists in the form of multiple socially homogeneous communities. In this case a socially homogeneous unit – one’s in-group – can punish its members who cheat outsiders and in doing so secure inter-group exchange. As Greif also points out,
Endogenizing fractionalization 79

however, this mechanism (the ‘Community Responsibility System’) cannot work if the disparate communities are internally large or socially heterogeneous in composition. In this event, agents run into the problems associated with social heterogeneity mentioned above.

This is problematic because most of the gains available from trade lie outside small in-groups and, in many instances where multiple groups are present, they are neither small nor internally homogeneous. As I discuss below, this was particularly true of communities in pre-colonial Africa. How, then, can socially distant agents capture the gains from widespread trade in this environment?

The failure of multilateral punishment in such situations does not mean that agents cannot bilaterally punish those who cheat them. In this case only the cheated party needs to know the identity of the cheater to prevent future trade. However, because cheaters only lose future trading opportunities with a single individual, bilateral punishment is not as effective in creating cooperation as multilateral punishment is when it is functional. Nevertheless, bilateral punishment can secure some, albeit a limited, degree of cooperation. When individuals bilaterally punish cheaters, agents who are sufficiently patient cooperate and those who are not cheat. Since the ability of *ex post* punishment to deter potential cheaters is limited when populous groups of socially distant individuals are involved, individuals cannot have as much faith *ex ante* that the outsiders they interact with will behave cooperatively. To overcome this uncertainty, individuals require a way of ensuring that they can avoid being cheated by impatient outsiders.

To do this they need to establish *ex ante* whether or not the outsiders they would like to trade with are ‘cheaters’ or ‘cooperators.’ In other words, they need a means of screening outsiders. Cheaters are agents with high discount rates, prone to defraud others in exchange. Cooperators in contrast have low discount rates and behave peacefully in trade. This binary division is of course artificial. In reality individuals lie somewhere along a spectrum of honesty with pure cheaters at one end and pure cooperators at the other. But, the general division is valid; there are those who tend to cheat and must be filtered out to avoid the losses associated with interacting with them, and cooperators who are profitable to interact with because they generate mutual gains from trade.

Screening, unlike multilateral punishment, does not depend on population size or social homogeneity to work. Successful screening does, however, require two things: easily observable attributes or activities – signals – that individuals may adopt or undertake to indicate their credibility to outsiders; and signals with an appropriate cost structure – namely signals that are cheap for cooperative types to send but expensive for cheaters to send – to effectively convey their credibility to outside potential trading partners they wish to exchange with. In other words, the cost function for sending some signal must satisfy the ‘single-crossing property’.
Signals that shrink the social distance between two agents satisfy both of these conditions. By creating some degree of homogeneity with respect to the outsider he wants to trade with, an individual signals his credibility, which in turn enables some level of exchange. If you and I are socially distant and I desire to trade with you, I can adopt certain customs or practices that belong to you and in doing so signal my trustworthiness in trade. I can use social distance-reducing signals to enable exchange between us. For instance, I might learn your language, join your religion, engage in your customs for settling disputes, or adopt some other practice of yours that reduces the social distance between us.\(^2\) These attributes are all easily observable. More importantly, however, it costs cheaters more to create homogeneity with outsiders over such dimensions than it costs cooperators.

The reason for this is straightforward. An individual’s payoff from reducing the social distance between himself and an outsider is long term. In other words, the costs of investing in ‘homogeneity capital’ with an outsider are only recouped through repeated play over time. Cheaters, however, have higher discount rates than cooperators. This is in fact why they cheat. Because they discount the gains from future exchange more heavily than cooperators, cheaters find it relatively more costly to invest in creating some degree of homogeneity with an outsider, the value of which will only be recouped sometime down the road. Following this logic, the more impatient the cheater, the more costly he finds the investment.\(^3\) If the cost of creating some degree of homogeneity is high enough (specifically, if this cost is greater than the one-period payoff from cheating), cheaters will not do so. Only cooperators will adopt this degree of homogeneity, so this signal can be used successfully to determine a sender’s credibility. If an individual observes a certain degree of homogeneity with an outsider, he knows that he is a cooperator, so he trades with him. If he does not observe this degree of homogeneity, he knows that the outsider is a cheater and so refrains from trading with him.\(^4\) In equilibrium this individual exhausts the gains available from trading with cooperative outsiders, while avoiding the costs of trading with cheaters.

Individuals use behaviors that reduce social distance to enable trade with outsiders instead of other costly behaviors for a number of reasons. The first is suggested by a repeated finding in the psychology and sociology literatures.

\(^2\) In the anthropological literature, Wobst (1977) introduced the idea that individuals use stylistic behaviors (appearance, symbols, seals, etc.) to manipulate their social distance from others.

\(^3\) See Posner (1998) for an excellent analysis of signaling in the context of separating cheaters from cooperators.

\(^4\) Williamson’s (1983) seminal paper on credible commitments and hostage taking similarly uses of the notion of costly investment for ensuring cooperation but in the context of firms and advertising. Williamson argues that in light of asymmetric information, which invites cheating, honest firms invest in advertisement and label promotion as a means of credibly assuring their customers they will behave honestly.
Individuals tend to be attracted by and place more trust in others who are like them in some way (see for instance, Kramer and Brewer, 1984; Cohen, 1977; Kandel, 1978; Verbrugge, 1977; Byrne, 1971; Newcomb, 1961; Hoffman and Maier, 1966; Lott and Lott, 1965). This is often so even when the dimension of homogeneity involved seems insignificant. For instance, if you are a fan of a certain football team and are wearing a hat with the team’s insignia and you find yourself in an elevator with a fellow wearing a t-shirt of the same team, you are likely to have an affinity towards this person. One of you may even initiate conversation with the other. The fact that you share this commonality serves as the basis for some amount of goodwill towards one another. The same is frequently true of encountering an individual you learn attended the same college or university as you. Even though you never met this individual in your time at that institution and she remains a stranger to you now, the fact that you share this dimension often creates a feeling of amicability towards one another.

In their well-known experiment, Tajfel et al. (1971) found a similar result. A population was divided along the seemingly arbitrary and insignificant dimension of homogeneity: taste in art. Specifically, participants were separated by their preference for paintings by Paul Klee or Wassily Kandinsky. Members of the ‘Klee group’ demonstrated a particular affinity for one another and likewise with members of the ‘Kandinsky group’ – an affinity that did not extend to those who differed with respect to this dimension of homogeneity.

Signals of credibility predicated on adopting degrees of homogeneity are therefore focal, and signal receivers look for them in evaluating the credibility of outsiders. In fact, in certain environments, specific degrees of homogeneity – i.e., commonalities over particular dimensions – are privileged over others and thus focal for cooperation, while others are not. At a football game, for instance, the team one supports may be focal and more important in shaping individuals’ beliefs about others that they do not know than other dimensions of potential commonality, for example language, that might be considered important in other environments.

As I discuss in Section 3, certain practices/customs achieved the privileged status of focal signals in pre-colonial Africa. In contrast to signals that reduce social distance, many other costly behaviors are not focal in this way. Thus, although their cost structure may permit them to function as signals in principle, because signal receivers do not generally recognize them as signals of credibility in practice, they do not function effectively as signals despite this.

Second, other things being equal, signal receivers prefer signals that appeal to their preferences. Again, perhaps for reasons psychological, perhaps evolutionary, individuals tend to derive pleasure from encountering others who share their interests, practices, tastes, style, etc. For example, individuals who believe in a particular religion are typically pleased to see others engaging in their religion. Economists like encountering people who have the same interests in economics as they do, etc. Signaling credibility to outsiders by adopting
such dimensions of homogeneity therefore creates a kind bonus of the signaling activity for receivers who, in addition to appreciating the signal for its ability to impart information about the sender’s type, also appreciate the signal in and of itself.

Third, some types of signals create a greater likelihood of repeated interactions than others. Repeated interactions are desirable because they build trust, reinforcing successful relationships and serving as the basis for further exchange. Degrees of homogeneity, such as religious practice, foster repeated interaction by their nature. Adopting the same religious practices as an outsider, for instance, may mean that you will encounter this person in church each week. Other costly behaviors do not have this quality.5

3. Signaling in pre-colonial Africa

This section applies the theoretical framework elaborated above to pre-colonial Africa. Specifically I examine how social distance-reducing signals enabled peaceful exchange between outsiders. While it is important to avoid romanticizing relations between Africans in the pre-colonial period, it is also important to recognize that ‘Long before the Europeans appeared on the scene’, Africans had established extensive domestic and ‘international trade, with developed systems of credit, insurance . . . [and] arbitration. Law and order were normally maintained and strangers honored their business obligations’ (Cohen, 1969: 6). Africans conducted much of their exchange on a long-distance basis and ‘[l]ong distance traders of all types banded together in caravans’ for this purpose (Hopkins, 1973: 62). Trade routes were in place throughout central Africa, sub-Saharan Africa, and other regions as well. Long-distance trade brought ‘about intensive social interaction between various ethnic groupings’. This trade involved ‘extensive credit arrangements often between total strangers from different tribes’ (Cohen, 1969: 6).

To make such exchange possible, agents adopted the customs and practices of the outsiders they wanted to trade with to signal their credibility. In this way ‘commercial interaction was an important element creating homogeneity’ among pre-colonial Africans (Thornton, 1995: 194). For instance, this practice along ‘trade-routes helps to explain the great cultural similarities between different peoples of the savanna south of the equatorial forest’ (Vansina,

5 A fascinating literature discusses the effectiveness of physical cues such as smiling, sweating and laughter, which are difficult to successfully fake, in signaling information about a party’s credibility/intentions to others (see for instance, Scharlemann et al., 2001; Ockenfels and Selten, 2000). In the discussion above I have emphasized the importance of signals that have the appropriate cost structure in enabling outsiders to determine an individual’s type. However, it may be that it is not the costliness of these signals that matters so much as it is the conversations they generate, particularly through repeated play. Such conversations put outsiders in a position to evaluate one another’s physical cues and in turn their credibility.
I consider three specific pre-colonial arrangements that served as focal signals enabling peaceful widespread trade in pre-colonial Africa: the individual’s relationship to authority and community, land usage, and religious practice/association.

### 3.1 Pre-colonial relationships between individual, authority, and community

Although forms of governance differed over time and across regions within pre-colonial Africa, stateless societies populated much of the continent. As much as 25% of West Africa, for example, was entirely stateless on the eve of colonization. The Tiv system of Nigeria extended to nearly a million people (Curtin, 1995a: 71), the Nuer system of the southern Sudan extended to over more than a quarter million people, and the Bedouin Arabs throughout Northern Africa all more or less functioned effectively on this basis too (Bohannan, 1968: 172). Not only were communities from the Tiv and Ibo of Nigeria to groups in Northern Uganda, Karamoja, and the East African Rift Valley stateless, they were largely heterogeneous in composition as well.

An even larger number of regions were quasi-stateless in the sense that formal political authority was so weak as to constitute virtually stateless social orders. Informal leaders often ‘headed’ society in both stateless and quasi-stateless regions. These informal leaders were nothing more than elders or other community members in high standing who might direct or advise community members when the situation called for it.

Under stateless or nearly stateless conditions, individuals subjugated themselves and expressed allegiance to leaders voluntarily. One could come and go between communities as pleased without fear of formal authority looming. Of course, many chose to remain affiliated with a particular community, ‘led’ by a particular leader, for long periods of time. With membership in such communities came the advantages that membership in a specific community confers. For instance, trade possibilities, assistance in times of need, and protection in the event that one was attacked were frequently linked to community association.

Although leaving the community brought no formal sanctions against the individual, informal social sanctions like ostracism or refused re-admittance were possibilities. Nevertheless, affiliation/membership in these societies remained voluntary in nature. Individuals sometimes gave gifts to leaders as a sign of good faith, but, like association with the leader and community itself, gifts were rarely mandatory.

Because gift giving and subjugation to community leaders constituted voluntary social activities, they effectively signaled the credibility of approaching outsiders who engaged in them to existing community members. An approaching agent’s *self-proclaimed* allegiance to a community’s informal leader indicated to existing members that he had an investment in this community he did not want to
lose. In this way social distance-reducing signals enabled approaching outsiders to peacefully exchange with the community’s members.

Pre-colonial African societies frequently resolved disputes within the lineages that comprised them and sought the advice of their informal leaders when necessary. For instance, in the stateless Lugbara society, ‘rainmakers’ – informal religious leaders – arbitrated conflicts (Middleton, 1971: 10). Of course the settlements arrived at were not formally binding. Like voluntary allegiance to a group’s leader, if an approaching outsider voluntarily agreed to settle his disputes informally in accustom with a community’s specific dispute rituals/practices, he signaled his credibility to the community’s members, which enabled exchange.

For example, various forms of oath taking, such as that employed by the stateless Kikuyu society in Kenya (Curtin, 1995b: 519), were often involved as part of the dispute settlement procedures that outsiders were required to submit to for exchange. Some communities confiscated individuals’ goods or imprisoned them for refusing to abide by settlement decisions. However, in many cases, if an individual ‘chose to ignore a ruling given by the chief, he could do so with impunity; but if public opinion was behind the chief’s decision, he might lose the privileges’ (Howell, 1968: 192) of membership in that community, including exchange. Taken together, the effectiveness of signaling, based on voluntary submission to community leaders and voluntary agreement to modes of arbitration, ensured that ‘Foreigners who settled among [new communities] were easily absorbed into local society’ (Colson, 1969: 29).

Elsewhere in pre-colonial Africa, where there existed more formal governing authorities, the state remained relatively weak, as characterized by the rarity of state adjudication and legislation. Consequently, customs and practices like oath taking and leader arbitration that served as important signals of credibility, enabling exchange in stateless and quasi-stateless societies, signaled credibility and made trade possible in more formal social orders as well.

This is particularly true of individuals’ voluntary subjugation to community authorities. Indeed, throughout pre-colonial Africa, ‘far from there being a single ‘tribal’ identity, most Africans moved in and out of multiple identities, defining themselves at one moment as subject to this chief, at another moment as a member of that cult, at another moment as part of this clan, and at another moment as an initiate in that professional guild. These overlapping networks of association and exchange extended over wide areas’ (Ranger, 1985: 248).

3.2 The pre-colonial structure of property arrangements
The structure of property arrangements in pre-colonial Africa also created an important means of signaling credibility to outsiders that promoted widespread exchange. Property arrangements were frequently structured in terms of land usage, such that the group of individuals using a piece of land was its primary ‘overseer’. The land-using community did not own the land, in the sense that it could sell the property to others, but it did often exercise some control over who
may use the land it was currently occupying. In some areas, informal community leaders advised group members as to what land parcels were unoccupied or might be good to work. In other regions, ‘Earth Priests’ – community leaders representing a link to the historical first user of the land – did not have the power to allocate land, but frequently performed this role as well.

In addition to this function, ‘Earth Priests’ usually had the charge of ensuring that members of the community followed certain ritual customs and taboos related to the believed mystical qualities of the land. Although the land-using community did not exert formal rights of exclusion over the property they used, outsiders who desired to enter the community frequently made gifts to the Earth Priests ‘as an expression of goodwill’ (Colson, 1969: 54). To gain access to the group, outsiders also typically agreed to respect the ritual customs and taboos established by the Earth Priests. An outsider’s voluntary agreement to engage in these practices signaled his credibility to existing community members, who would then welcome the outsider to join them. By joining the land-using community, an outsider was offered the privileges typically associated with membership, among which included exchange.

While many were sedentary agriculturalists, some communities of land users migrated frequently, establishing usage over new areas of land and leaving old areas behind. This gave new outsiders the opportunity to interact with the group as it traveled. By adopting the group’s practices, such as giving special gifts to the Earth Priest and agreeing to respect its ritual taboos, new outsiders signaled their credibility and joined the migrating community. By joining the group, these outsiders were given the possibility of trading with existing group members.

Migration also gave existing community members the option of staying behind if they so desired. Group composition was constantly changing as a result of new and old outsiders’ ability to signal trustworthiness with the aim of cooperation. As a result of this practice, throughout pre-colonial Africa, heterogeneous ‘people were linked to land through their membership in groups’ (Chanock, 1985: 231) and vice versa, making exchange possible with a wide variety of individuals.

3.3 Pre-colonial religious practices and associations

Finally, religious practice and association also acted as focal signals enabling exchange among pre-colonial Africans. By participating in the religious practices and beliefs of an individual, completely converting to his religion, or joining his religious association, approaching outsiders signaled their credibility to this individual, enabling trade. Indeed, because of this of practice, long-distance traders in pre-colonial Africa were frequently bound together by religion.

For instance, Islam ‘helped maintain the identity of members ... and enabled traders to recognize, and hence to deal readily with each other; and it provided moral and ritual sanctions to enforce a code of conduct which made trust and credit possible’ (Hopkins, 1973: 64). Even up through the twentieth century, ‘Indigenous traders become Moslems in order to partake in the moral community
of traders. In both Ibadan and Freetown, nearly half the population is Christian. Yet in both cities all the butchers without any exception have converted to Islam, because only in this way can they participate in the chain of trade' (Cohen, 1971: 278). Other pre-colonial Africans converted to Christianity to signal credibility to the European outsiders they desired to trade with.

Secret societies in pre-colonial Africa – typically quasi-religious organizations – often performed both religious and judiciary functions in stateless and quasi-stateless societies (Curtin, 1995a: 72). In many places they were ‘a central part of the process of . . . handling conflicts’ among individuals (Chanock, 1985: 86). By joining these religious associations, agents powerfully signaled their credibility to existing members.

For some secret societies like the Ekpe, payment was sometimes required and membership restricted. In other cases ‘cult membership was open to any who wished to join’ (Colson, 1969: 59). In signaling credibility to outsiders, secret societies ‘pull[ed] men from the domestic routine to meet with nonrelations in a wider social world, with its own . . . ceremony and ritual’ (Ottenberg, 1971: 154).

Participation in religious brotherhoods like the turuq acted as signals in a similar fashion. Again, by joining these brotherhoods, approaching outsiders conveyed their credibility to existing members, making possible exchange. As a result of the signaling made possible by religious practices and associations, ‘[a]mong the ethnic groups in the Cross River and Niger Delta area, [for instance], where the societies were decentralized, order was maintained by a balance of understanding among the associations, overlapping memberships, and respect for the rules of public conduct generally recognized in the community’ (Awe, 1999: 7).

4. Bad institutions and the breakdown of exchange

Critical to the success of the signaling mechanism described above is a ‘good’ institutional framework. Because the ability of individuals to interact with outsiders depends crucially on the fact that their signals are clear and reliable, a ‘good’ institutional environment, in addition to not being ‘extractive’ (as emphasized by Acemoglu et al., 2001), also does not interfere with agent signals. In contrast, a ‘bad’ institutional environment not only creates a high ‘risk of expropriation’, but also generates noise that causes the signaling process described above to break down.

Absent institutional interference, traders willingly adopt the practices of outsiders they desire to interact with because they believe that the gains from peaceful exchange outweigh the cost of investment in the signal. In other words,
in a noise-free environment, all signals sent are consistent with the sender’s underlying type as a cooperator (a low discount rate type) rather than a cheater (a high discount rate type) who would not send the signal because of its cost. Signals therefore effectively communicate information about the signaling party’s credibility.

Signals sent against a ‘bad’ institutional backdrop, however, are not necessarily consistent with the sender’s underlying type. Institutional interference creates false signals regarding the trustworthiness of particular individuals in trade. Consider the following example involving two socially distant traders, Trader 1 and Trader 2.

Absent institutional interference, Trader 1 can trust that, if Trader 2 joins his religious society, he must have a sufficiently low discount rate and will therefore cooperate in exchange, but, if Trader 2 does not send this signal, Trader 1 knows that he has a high discount rate and will cheat. However, if authorities formally prohibit religious societies, Trader 1 cannot be sure whether Trader 2’s failure to join his religious society reflects Trader 2’s underlying type as a high discount agent or is merely the outcome of the state’s prohibition.

Similarly, imagine that formal authorities made it mandatory for all citizens to pay allegiance to community leaders and to submit to specific, codified dispute settlement procedures. Again, Trader 1 would have no way of knowing whether the fact that Trader 2 did these things reflected his underlying type as a low discount agent or simply reflected the state’s fiat that he do so.

In this environment, traders are unsure how much of the credibility signaled (or not signaled) by approaching agents is genuine and how much is artificial. Agents’ genuine signals of credibility to one another interact with and become indistinguishable from the artificial signals generated by the new institutions. In other words, agents face a signal extraction problem. In this way, institutional interference creates noise in the signaling mechanism that reduces its informational efficiency. As a result of this noise, which prevents individuals from being able to determine the credibility of outside potential exchange partners, many transactions that would have taken place do not.

Without clear signals to gauge the credibility of outsiders, agents refrain from interacting with them. Instead they interact only with those they know very well—their ‘in-group’—since they do not need to rely on signals to assess the reliability of these exchange partners. They may even create smaller sub-groups within their broader in-group as a means of eliminating the risk of interacting with those who are even remotely unknown. This reversion to in-group interaction and the creation of new and smaller sub-groups within in-groups polarizes agents and increases fractionalization. Because individuals cannot capture the gains from widespread trade, low economic growth results.

There are two reasons why individuals could not simply shift to other kinds of signals in the face of the problem created by colonial institutions described above. First, colonial-created prohibitions incapacitated the focal signals that
enabled widespread trade. In other words, it is not merely that colonization
disabled some potential signals for facilitating exchange; colonization disabled
the very signals that individuals looked to when evaluating the credibility of
outsiders for trade. For new signals to be effective, they would first need to be
established as focal. But the evolution of focal signals takes considerable time.
In Africa, at least, this evolution was the product of a lengthy process, which
only eventually resulted in the widespread coordination of a large number of
socially disparate individuals based on certain practices and customs that would
be recognized as signals of credibility. There is no reason to think that it would
be immediately obvious to a large population of socially diverse individuals what
new signals they should coordinate on to make cooperation possible in the face
of the colonial-created signal extraction problem.

Second, it stands to reason that in the absence of state-created signal noise,
agents are employing the least-cost signals possible that will enable exchange
with outsiders. For instance, if, in the absence of colonial law, individuals use
religious practice/association to enable trade, it is reasonable to believe that
this is the least-expensive signal that is still costly enough to filter out cheaters.
If not, agents would use other forms of signaling in its place. If this is true
and colonial law disables this means of signaling, agents must use relatively
more costly signals to make possible exchange. By raising the cost of enabling
widespread trade, institutional noise reduces the number of exchanges between
socially diverse individuals, increasing fractionalization.

5. Signal noise, fractionalization and colonial institutions

This section considers how three colonial institutions: the Native Authority,
property law, and religious policy, diminished trade between socially diverse
agents by removing their power to convey credibility through the social distance-
reducing signals discussed in Section 3. While different colonial powers in Africa
ruled differently and ruling strategies changed over time, their essential features
are sufficiently similar to generalize their outcome on the ability of Africans to
engage in widespread trade.

For instance, although the varying powers emphasized ‘indirect’ vs. ‘direct’
rule over time, ‘the tendency . . . has been . . . to exaggerate differences rather than
similarities between colonial policies’ when in fact ‘there was little difference
in their overall effects’ (Kiwanuka, 1970: 1, 4). The bulk of my discussion
comes from colonial policy under ‘indirect’ rule, whereby colonizers attempted
to govern through indigenous institutions and custom as much as possible.
However, instances of ‘direct’ rule do not significantly alter my analysis.

5.1 Noisy signals regarding relationships to authority under colonialism

In the pre-colonial period heterogeneous individuals cooperated by signaling
credibility through their relationship to community ‘leaders.’ However, by
significantly adulterating these relations, the colonial institution of the Native Authority rendered these signals ineffective. As a result, conflict emerged between agents, polarizing them from one another.

Most colonial powers created the institution of the Native Authority or Native Administration under the auspices of ‘indirect’ rule to govern Africans through native rulers in each community. In order to administrate effectively, the colonial powers were interested in establishing communities with clearly defined rulers. Indeed, their preconceptions about ‘traditional Africa’ included the notion that rulers were at the center of every society. ‘More than just creating a tribal hierarchy where none had existed previously, this often involved working through a mishmash of ethnic affiliations to create ‘purer’ and clearer tribal identities as the bases for tribal authorities’ (Mamdani, 1996: 81).

The first step in establishing such rulers was identifying who in each case was to rule. Where states previously existed, colonial administrators endowed pre-existing chiefs with colonial-backed authority as the rulers of their communities. Over the large regions where no formal rulers previously existed, colonial authorities found identifying appropriate rulers considerably more difficult. To deal with this problem, they often appointed and installed previously informal leaders or community elders as *formal* rulers in a region.

Elsewhere, where colonial authorities could not identify leaders, they simply selected rulers arbitrarily to govern the community (Roberts and Mann, 1991: 21). Owing to the migratory nature of some societies, colonizers could not always even identify formal communities. In these instances, the colonists created new ethnicities and agglomerated them under the rule of a colonial selected native authority, as was the case with the colonial-created Ndebele ethnic identity in Zimbabwe (Mamdani, 1996: 81). Thus, colonial-imposed ‘chiefs were not . . . those who would have been selected according to [pre-colonial] customary procedures’ (Crowder, 1970: 213).

In contrast to pre-colonial arrangements, the Native Authority required community members to follow the dictates of the ruler imposed over them. For instance, S.25 of the 1927 Native Administration Act in South Africa was typical in this regard of colonial acts concerning the institution of the Native Authority. This administrative act empowered chiefs to ‘rule all natives by decree’ (Mamdani, 1996: 71). Colonial policy also imbued chiefs with the power of taxation and in places like Uganda made it illegal for subjects to ‘insult a chief’ or ‘refuse to pay homage to a chief’ (Grier, 1987: 38). Under these conditions an individual’s allegiance to the community leader could not effectively signal information regarding his credibility. The Native Authority, which made leader allegiance mandatory, had the effect of making submission to leaders ambiguous in terms of its ability to signal an agent’s credibility to others.

The rulers installed by colonial administrators were also given new powers as the formal arbiters of disputes. ‘Not only did the chief have the power to pass rules governing persons under his domain, he also executed all laws and was
the administrator in ‘his’ area, in which he settled all disputes. The authority of the [colonial] chief thus fused in a single person all moments of power: judicial, executive, and administrative’ (Mamdani, 1996: 23). For instance, the 1891 Natal Native Code endowed chiefs not only with the authority to command ‘respect and obedience’ (S.57) from his subjects, but also ‘to try all civil cases (divorces excepted) between natives’ (S.49), to impose fines and to arrest.6

In this way the Native Authority formalized, coercively enforced, and unified functions previously dispersed among voluntary-based informal associations like age sets, clans, and families, in the hands of a single authority. Although the colonial powers viewed this as simply an extension of existing custom, in actuality nothing could be further from the case. As noted in Section 3, in the pre-colonial period, decisions to settle disputes via the advice of the community elder were voluntary. ‘There was no automatic jurisdiction. When people wanted or needed an arbitrator they looked for and found one’ (Chanock, 1985: 34). Even in communities that had more formal governance structures a chief ‘could not enforce his rule on senior headman and the latter could not exact obedience from villagers’ (Turner, 1957: 14–16).

The colonial institution of the Native Authority changed all of this by endowing the appointed chief with the power to adjudicate all native disputes via native courts and enforce court decisions. In this environment, the fact that an individual agreed to have his disputes settled by the community leader and followed through on the leader’s decisions did not clearly signal credibility as it once did. Again, existing community members could not determine whether such behavior on the part of outsiders reflected their underlying type as a credible exchange partner or whether it was merely the result of the law compelling them to do so.

Conflict between individuals as a result of this colonial-created signal noise began to manifest itself during the colonial period. This conflict took several forms. First, colonial-imposed chiefs with new powers to coerce communities caused numerous conflicts throughout Africa. For instance, ‘A major cause of the riots’ in Eastern Nigeria ‘had been the imposition…of the warrant chiefs and their courts, judicial institutions of a type which had no roots in local society’ (Chanock, 1985: 26) and consequently created a signal extraction problem for individuals. Anti-chief riots emerged elsewhere in places like Sierra Leone in 1955–57 as well (Mamdani, 1996: 106).

In addition to this, conflict appeared in the form of rapidly increasing legal battles. Indeed, the courts became ‘more and more the sole mechanism through which the social system is maintained…the Nuer [for example] are pressing…claims for the fulfillment of obligations which were not in the past normally the subjects of dispute at all.’ In his study of the Barotse, Gluckman commented that ‘The Lozi, like all Africans, appear to be very litigious. Almost

Endogenizing fractionalization

every Lozi of middle age can recount dispute after dispute in which he as been involved’ (Gluckman, 1955: 21).

In the colonial-created environment of noisy signals, ‘obligations of social existence are more and more maintained only by direct action through the courts’, symptomatic of a ‘wider process of social disintegration’ among Africans (Howell, 1954: 231, 236). For example, colonization gave rise to ‘escalating numbers of civil conflicts over marriage, divorce, inheritance, and succession initiated by litigants professing different customs’ (Roberts and Mann, 1991: 21). In short, signal noise generated by the Native Authority stifled peaceful interaction among heterogeneous agents because ‘the role expectations’ played ‘in the past no longer provide[d] a basis for mutual understanding and adjustment but rather bec[a]me a source of conflict’ (Epstein, 1981: 351).

In this environment, individuals became suspicious and extremely cautious in their dealings, even with their kin. They were of course even more reluctant to engage with outsiders. The Native Authority infected custom-based signals with ‘uncertainty’ and ‘ambiguity’, and as a result ‘trivial incidents often led to violent responses’ (Chanock, 1985: 228). Indeed, ‘The more custom was enforced’ by colonial-imposed rulers, ‘the more the tribe was restructured and conserved as a more or less self-contained community . . . as it had never been before.’ Community polarization piqued in the form of various nationalist movements throughout Africa in the period between World War I and World War II (Mamdani, 1996: 51, 102).

5.2 Noisy signals regarding property arrangements under colonialism

The colonial institution of property ownership also incapacitated important signals between heterogeneous agents. In pre-colonial Africa, association with and usage of a particular piece of land created ‘membership’ in a certain community of existing land users, and this membership in turn made possible exchange. However, ‘key’ to colonial policy, which ultimately ‘defined a world of the customary from which there was no escape . . . was the definition of land as a customary possession’ (Mamdani, 1996: 21). In this way the colonial institution of property introduced a previously unknown rigidity into individuals’ relationships to land.

The European colonizers brought with them ‘European concepts of legal tenure’, which they assumed ‘must exist in Africa as in Europe’ (Colson, 1971: 196). So, much as they sought to establish formal rulers for the African communities over which they governed, they also attempted to install the institution of proprietary ownership among African communities. However, as discussed in Section 3, in the overwhelming number of instances in pre-colonial Africa, proprietary owners did not exist.

Certain groups of individuals might be the predominant users of some region of land, but this did not constitute ‘ownership.’ Furthermore, groups frequently migrated over time, leaving certain regions of land and entering others. Where
no private person appeared to hold such rights over a given area, [colonizers] assumed that the rights must vest in the political unit whose members used the region’ (Colson, 1971: 196). Colonial administrations thus vested the formal right to allocate land in Earth Priests, where they could find them, and, failing this, they gave these powers to the colonial-installed chiefs.

Under a land policy that empowered rulers to allocate land usage, an outsider’s use of the same area of land need not reflect the fact that the land-using community had admitted him under the belief he was credible. It might only reflect the ruler’s command. In contrast to pre-colonial times, under colonization, individuals ‘could acquire rights in land without also acquiring citizenship in a local community’, and rulers could ‘convey their rights to a stranger who was not necessarily acceptable to the community’ (Colson, 1971: 211). In other words, outsiders could not be screened according to their credibility through the use of signals relating to a group’s land practices. As such, high discount agents who were likely to cheat could gain membership in the community. This created a great deal of uncertainty in exchanging with outsiders and encouraged individuals to resort to exchange among family members or the few that they knew extremely well.

Often times colonial land policy prevented individuals who had been allocated a particular piece of land from disassociating themselves from this land. For instance, the Native Code of 1891, which applied to the Natal and Cape regions of Africa, decreed that the ruler could ‘create and define pass areas within which natives may be required to carry passes’ and could ‘prescribe regulations for the control and prohibition of the movement of natives into, within, or from such areas (S.28)’. This was the case, for example, in Lugbaraland beginning in 1914. In many instances ‘internal migrations were controlled’ or prohibited by colonial law as well (Gann and Duignan, 1978: 366). For instance, the Calendon Code of the South African Cape in 1908 made it ‘compulsory for a Khoisan to have a fixed and registered place of residence and carry a pass when moving within or between districts of a colony’ (Hindson, 1987: 15–16).

In pre-colonial Africa, where individuals could come and go as they pleased, it was possible for them to be ‘members’ of multiple communities and as such to exchange with a wide range of different individuals. Furthermore, agents’ ability to exit voluntarily social groups strongly signaled their credibility to existing group members where they stayed. By legally requiring individuals to remain attached to their ruler-allocated areas of land, colonial land policy created noise in the signals used to convey credibility to outsiders. An individual’s connection to the same land as other members of a community need not constitute a genuine signal of his credibility. His stay on the land could just as easily be the result of the law that required him to remain there. Similarly, colonial policy in restricting migration, restricted individuals’ ability to signal credibility to members of other

communities. This of course diminished their ability to interact peacefully with outsiders as well.

In other places, colonial policy explicitly prohibited rulers from allocating land to outsiders. For instance, in 1903 in Basutoland colonial law forbade the provision of land to non-Basotho people. Elsewhere, colonial policy restricted the kind of land access that outsiders could have. For example, the Native Authority Rules of 1948 in Nigeria required outsiders to obtain a special license in order to farm (Mamdani, 1996: 140). In some places, government would not recognize private individuals’ agreements regarding land. Like the other features of colonial land policy, these restrictions helped destroy the signaling capacity of engaging in the land rituals and practices of outsiders as a means of making exchange possible.

The breakdown of individuals’ ability to interact peacefully as a result of this signal noise manifested itself in the sharp rise in property disputes between Africans under colonialism. Instead of property usage creating membership of a group and cooperation, it became a primary point of conflict. ‘Communities sought to extend the boundaries of their holdings and sued one another over land that both had formerly ignored.’ ‘Before the end of the Belgian Congo regime’, for example, ‘local courts were overwhelmed with land suits as community sued community and private persons claimed the right to share in the profits realized from a communal estate’ (Colson, 1971: 207). Thus, the ‘effect of the new property’ institution was ‘corrosive’ (Chanock, 1985: 37) on cooperation, as it generated constant conflict between individuals, resulting in increased polarization that inhibited widespread exchange.8

5.3 Noisy signals regarding religious practice and association under colonialism

Finally, colonial policy regarding native religious practice and organizations incapacitated the signals between diverse individuals that made possible widespread exchange. As discussed in Section 3, in pre-colonial Africa traders frequently used ‘their religious reputation to pass back and forth’ through different areas (Curtin, 1995a: 174). Their association with the religions of various outsiders signaled their credibility to these individuals, enabling trade.

Colonial powers in most places worked closely with Christian missionaries to curtail what they considered offensive native religious practices. Although

8 Bates, Greif and Singh (2002) point out that under situations of weak third-party enforcement, rising property values lead to increased conflict over this property. Their argument is important in explaining increases in property-related conflict under statelessness and may be relevant for explaining part of the increase in property-related conflicts in colonial Africa in areas where colonizers did not introduce third-party enforcement. However, its importance in accounting for the surge in property conflicts under colonial institutions in many other areas is likely small. The simple reason being that colonial rule typically substituted strong third-party enforcement for the weak third-party enforcement that existed in the precolonial period.
the General Act of the Berlin Conference in 1885 guaranteed locals religious freedom, in practice colonial authorities tolerated native religious customs only so long as they did not see them as threatening to their power. In most instances, ‘From the point of view of the colonial state, rituals, religious activities, and large social gatherings were sites of potential ‘subversion’, ‘and so needed to be extinguished (Likaka, 1995: 203). For instance, in Nkole, ‘shrines were torn down, cult objects confiscated and cult leaders discredited, fined, or imprisoned;’ indeed, ‘the logical and institutional basis of [individuals’] mystical relations to king and state diluted’ (Stenning, 1965: 266).

Elsewhere colonial law explicitly prohibited certain religious practices and associations. For example, colonial law in Igboland contained an ‘Unlawful Societies Ordinance’ (Kalu, 1977: 79). Ordinances like these were designed to crack down on secret societies, which were formerly used as a means of signaling credibility that enabled exchange. In other cases colonial law directly outlawed religious practices like witchcraft and other ‘black arts’ and made it criminal to accuse others of such practices as well (Gann and Duignan, 1978: 235). Similarly, colonial authorities often forbade ordeals that pre-colonial Africans sometimes used to determine the guilt or innocence of accused parties. In Zaire even ‘night dances and most rituals like Matamba had been prohibited’ (Likaka, 1995: 203).

These prohibitions had a negative effect on the capacity of individuals to signal credibility to one another. Colonial criminalization of many religious practices and certain quasi-religious associations like cults and secret societies effectively eliminated these practices and organizations as potential signals of credibility to outsiders. In pre-colonial Africa, the absence of an individual’s association with religious practices/organizations was just as important as his association with such practices/organizations in conveying information about him. However, under colonial prohibition the fact that an individual did not practice witchcraft or belong to a particular cult, for example, could not signal information about his credibility. And as a result of colonial prohibitions like those on ordeals, ‘conflicts were sharpened’ (Chanock, 1985: 102) among Africans, curtailing the possibility of extensive exchange.

6. Conclusion

My analysis has important implications for understanding the connection between institutions and fractionalization. First, contra the position advanced by the fractionalization approach, my argument suggests that the presence of socially diverse individuals per se is not the cause of low growth. In pre-colonial Africa, where a large number of diverse individuals existed, agents were able to realize many of the gains from widespread trade by voluntarily adopting the practices, customs, and institutions of outsiders. Specifically, individuals’ relationships to authority and community, the structure of land usage, and religious associations performed this function. If agents can and
have engaged in widespread trade where socially heterogeneous individuals coexist, the mere presence of such diversity cannot explain poor development. Instead, my discussion emphasizes that what is important for economic growth is the ability of individuals to realize the gains from trade, regardless of their heterogeneity. Social distance-reducing signaling plays an important part in making this possible.

Second, my analysis suggests that destructive fractionalization is endogenous to the poor institutions that create it. In other words, the negative effect of fractionalization on economic growth is second order. Institutions determine the first-order effect. Although heterogeneity per se does not prevent agents from capturing the gains from exchange, heterogeneous agents are susceptible to the effects of ‘bad’ institutions that interfere with the signaling mechanism that enables them to trade. In Africa the colonial institutions of the Native Authority, property law, and religious policy did just this.

My analysis also offers potentially fruitful avenues for future research. One such undertaking might examine the extent and specific form of institutional interference with signals between heterogeneous agents in other struggling economies. Alternatively, it would be interesting to consider to what extent institutional interference with agent signaling is limited, or not limited, to foreign-imposed institutions, which may not be as well suited to particular local conditions as those imposed by domestic authorities. This application of my framework could also be used in exploring the effects of foreign-directed post-war reconstructions and foreign-directed post-socialist transitions on the ability of individuals in these places to engage in widespread exchange.

References


