Part II Institutions as Systems in Equilibria

Definitions direct analytic attention and confine the scope of an empirical study; conducting a study requires an analytical framework. An analytical framework is required to reveal the conditions under which a particular institution is effective in generating a particular behavior, to expose causal relationships, to generate predictions, and to evaluate arguments. An analytical framework is particularly important in studying institutions, because beliefs and norms are unobservable.[1]

Central to the analytical framework used here is classical game theory. Because the proof is in the pudding, this book contains five empirical studies that attest to its usefulness in studying institutions as defined here. These studies use an empirical method that combines detailed contextual knowledge of the situation and its history with explicit, context-specific modeling. This case study method uses contextual knowledge to develop a conjecture regarding the relevance of a particular intertransactional linkage and the related institution. It then uses an explicit (and in this work) game-theoretic model to evaluate this conjecture. This empirical method is employed in all the historical analyses presented in this book. It is elaborated on in Part IV.

The empirical studies presented here consider the institutional foundations of markets. Doing so departs from a long tradition in institutional analysis that goes back to Adam Smith and considers markets as primitives that need not be explained. According to this view, markets emerge spontaneously when and where there is an opportunity for profitable exchange. As O. Williamson put it, "in the beginning there were markets" (1975, p. 20).

The failure of markets to emerge, particularly in transition economies, revealed the fragility of this assertion. Markets do not necessarily spontaneously emerge in response to opportunities for profitable exchange. For exchange to transpire, institutions that protect property

[1] Indeed, as W. Powell and DiMaggio (1991, p. 2) note, promising institutional research that attempted to go beyond the institutions-as-rules and agency views—from Veblen and Commons in economics to Parsons and Selznick in sociology—"fell into disfavor, not because they asked the wrong questions, but because they provided answers that were either largely descriptive and historically specific or so abstract as to lack explanatory punch."

rights and provide contract enforcement must be in place. By determining who can exchange and what products can be exchanged, these institutions determine the scope and scale of the market.[2]

The next two chapters demonstrate the usefulness of the perspective developed here for empirically studying the institutional foundations of markets. Chapter 3 considers private-order institution for contract enforcement. Chapter 4 presents an institution that secured the property rights of alien traders. By dealing with distinct institutions, these analyses lend empirical support to the claims made in Chapter 2. In analytically studying institutions, it is imperative to examine intertransactional linkages and the interrelated institutional elements rendering this linkage effective and generating regularities of behavior.

Chapter 5 reflects more generally on the relationship between game theory and institutions as defined here. It seeks to gain insights about institutions by learning from the unrealistic assumptions that are required to make game theory a useful tool for studying behavior in strategic situations. The discussion first highlights what classical and learning game theory teach us about the role of institutionalized rules in providing the cognitive, informational, coordinative, and normative micro-foundations of behavior. It then proceeds by exploring the implications of this better understanding of rules and their relationships to beliefs about the appropriate use of game theory for institutional analysis.

When presenting and analyzing institutions, it can be cumbersome to adhere to the terminology presented in Chapter 2. It is easier to note, for example, that a community provides a network for information transmission than to say that linkages of information-sharing transactions within the community imply circulation of information. It is easier to note that members of a community shared the expectation that a cheater would be collectively punished than to say that the institutionalized beliefs for collective punishment prevailed within the community. Similarly, it can be easier to state that a particular (institutionalized) rule of behavior prevailed than to say that beliefs and norms motivated behavior corresponding to these rules. For ease of exposition, therefore, I often adopt the simpler rather than the more precise terminology.

---

[2] For related research, see Greif (1989, 1992, 1994b, 1997a, 2000, 2004b); Milgrom, et al. (1990); Stiglitz (1994); Greif and Kandel (1995); Aoki (2001); McMillan (2002); and Fafchamps (2004).

Chapter 3 Private-Order Contract Enforcement Institutions: The Maghribi Traders' Coalition

In premodern trade, merchants had to organize the supply of the services required for the handling of their goods abroad, because goods were sold abroad only after being shipped to their destination (De Roover 1965; Gras 1939). A merchant could either travel with his merchandise or hire overseas agents to handle his affairs abroad. Employing agents was efficient since it enabled merchants to avoid the time and risk associated with traveling and to diversify their sales across trade centers. Despite their efficiency, however, agency relations are not likely to be established unless supporting institutions are in place, because agents can act opportunistically and embezzle the merchants' goods.

This chapter examines the reputation-based economic institution —which can be referred to as a coalition—that enabled the Maghribi traders, a group of Jewish traders in the Mediterranean in the eleventh century, to deal with the contractual problem inherent in the merchant-agent transaction. In reputation-based institutions, future rewards or penalties in (auxiliary) economic or social transactions are made conditional on conduct in a central transaction. When effective, this intertransactional linkage enables an individual to credibly commit himself ex ante not to behave opportunistically ex post. In the case of agency relationships, the agent can commit to honesty and hence be trusted. Examining the operation of such institutions requires studying the institutional elements that create the intertransactional linkages and allow future utility to be conditioned on past conduct. In particular, such examination has to identify which transactions are linked to create sufficiently large sanctions or rewards, how information about past conduct is generated, and why inflicting sanctions or providing rewards is credible. (Appendix C presents the analytics of such institutions.)[3]

Two intertransactional linkages were central to the Maghribi traders. First, the agent-merchant transaction was linked with information-sharing transactions among the merchants. The resulting network provided merchants with the information required to evaluate the conduct of

[3] In the game-theoretic setting, a player's reputation is defined as a function from the history of the game to a probability distribution over his strategies.

faraway agents. It supported the institutionalized beliefs that opportunistic behavior is likely to be detected. Second, the agency transaction between each merchant and agent was linked with future transactions between that agent and every other merchant in the group. Every merchant in the group was expected to hire only member agents and never to hire an agent who had cheated another member.

A credible threat of collective punishment supported the beliefs that the short-run gain from cheating today was less than the long-run benefit of being honest. Because this situation was common knowledge, merchants perceived that the agents could not do better by cheating. A member agent acquired a reputation of being honest, and the merchants could trust him while a set of rules— known as merchants' law—defined what actions constituted appropriate conduct. [4]The Maghribis' code of conduct was a social norm, a rule that is neither promulgated by an official source, such as a court or a legislator, nor enforced by the threat of legal sanctions but is nevertheless regularly complied with (Posner 1997).

The organizational manifestation of this institution was an informal organization—a business network of members who belonged to the same ethnic and religious community. It was a reflection of and a means for creating an intertransactional link that changed the set of self-enforcing beliefs in merchant-agent transactions. It was the manifestation of the institutionalized beliefs that individuals with a particular social identity—those who belonged to the community—would share information and collectively punish a cheater. Indeed, the existence of this community and the personal familiarity and information flows within it facilitated the rise of the coalition. At the same time, the benefits of transacting with other community members were greater than those each trader could have realized by establishing agency relations based on a reputation mechanism outside the group. Each member was therefore motivated to maintain his communal affiliation, thereby perpetuating this social entity.

Analysis of the Maghribi traders is based on a historical source found in Fustat (Old Cairo) known in Hebrew as the *geniza* (depository). It contains about a thousand contracts, price lists, traders' letters, accounts, and other documents that reflect eleventh-century trade in the

[4] In the game theoretic setting, a player's reputation is defined as a function from the history of the game to a probability distribution over his strategies.

Muslim Mediterranean.[5] These documents were written by the Maghribi traders, who lived initially mainly in the western basin of the Mediterranean. (The *Maghrib* is the Arabic word for the Muslim world's *West*.) For religious reasons, these traders deposited every document written in Hebrew letters in the *geniza* of a synagogue in Fustat. Because they conducted their commercial correspondence in Judeo-Arabic—an Arabic dialect written in Hebrew letters—it is reasonable to conjecture that the documents found in the *geniza* contain a representative sample of their commercial correspondence (Goitein 1967, p. 149).[6]

In section 3.1, I provide information on the context and the Maghribi traders and describe their pattern of behavior in agency relationships and its purpose. Then in section 3.2, I discuss the commitment problem inherent in the merchant-agent transaction and evaluate the historically derived conjecture that a multilateral reputation mechanism mitigated the associated commitment problem. Section 3.3 models the commitment problem in order to determine whether the institution postulated here could have constituted an equilibrium and why. Section 3.4 uses the model to generate predictions that further substantiate the claim that agency relations were governed by a coalition. It discusses the role of the merchants' law in coordinating collective responses among the Maghribi traders. In section 3.5, I consider the implications of the analysis.

3.1 Commerce, Overseas Agents, and Efficiency

[5] For an introduction to the *geniza,* see Goitein (1967, intro.). The documents were purchased by various libraries. Documents referred to here are denoted by the library in which they are located and their registration number. When the reader is directed to specific lines within the document, the side (a or b) and the lines are also indicated. The list of abbreviations used here appears at the beginning of the book. Many of the documents have been published by Goitein, Gil, and others. For published, translated, or quoted documents, I cite the published source after the reference to the document. For example, TS xx.xxx, a, lines 24–25, Goitein (1967, p. 727) is a reference to document xx.xxx in the Taylor-Schechter collection, side a, lines 24–25, published in Goitein 1967 on page 727.

[6] On Judeo-Arabic, see Blau (1961, 1965). This chapter is based on Greif (1989, 1993). The analysis is based on about 250 documents contained in Greif (1985), Michael (1965), Gil (1983b), and Ben-Sasson (1991). These documents are the only ones available regarding trade between Egypt, Sicily, and Israel during the mid-eleventh century and the trade of Naharay ben Nissim, a Maghribi trader who lived in Fustat earlier that century.

The Maghribi traders were descendants of Jewish traders who left during the tenth century, the increasingly politically insecure surroundings of Baghdad and initially emigrated to Tunisia in North Africa —part of the Muslim West, the Maghrib—which was controlled by the Fatimid caliphate. Toward the end of the century, the capital of the Fatimid caliphate moved to Cairo. The Jewish traders that followed them from the Maghrib became known in Egypt as the Maghribi traders, the traders whose place of origin is the Maghrib.

The Maghribi traders were a minority, with a distinct social identity, within a much larger Jewish population. We do not know how many Maghribi traders operated during the eleventh century, but we do know that the number was not trivial: in 175 documents, 330 different names are mentioned.[7] Most of the Maghribi traders invested in merchandise worth several hundred to several thousand dinars—substantial sums at a time when the monthly expenses of a middle-class family in Fustat were two to three dinars.[8]

The *geniza* documents indicate that eleventh-century Mediterranean trade was free, private, and competitive. They show that there were few official restrictions fettered migration or the transfer of raw materials, finished goods, or money across the Mediterranean.[9] Within each trade center, commercial transactions were conducted competitively. But trade was characterized by uncertainty over such factors as prices, the duration of the ship's voyage (and whether the ship would reach its destination at all), the condition in which the goods would arrive, and the cost of storage.[10]

---

[7] These letters are all those available regarding the trade with Sicily and Israel during the mid-eleventh century and the trade of Naharay ben Nissim. They are a subset of the documents specified in the previous footnote.

[8] See Goitein (1967, p. 214ff..); Gil (1983b, 1: 200ff..); Greif (1985, pp. 73–6). For expenses, see Goitein (1967, p. 46) and Gil (1983a, p. 91).

[9] Customs, however, were imposed. See Gil (1983b, 1:257–8); Goitein (1967, pp. 29–35, 157, 187, 192 ff., 266–72); and A. Lewis (1951, pp. 183-224).

[10] See, for example, Dropsie 389, a, lines 4–5, b, lines 27–8, Gil (1983a, pp. 113–25). See also Goitein (1967, pp. 148ff, 200–01, 273 ff., 301); Stillman (1970, pp. 70-82ff); and Greif (1985, pp. 3, 69–78, 92; 1989, 1993).

To cope with the uncertainty and complexity of trade, the Maghribi traders operated through overseas agents. An overseas agent is anyone who supplies the services required for a commercial venture and shares the capital, profit, or both with a merchant located in a different trade center. (Henceforth in this chapter and Chapter 9 the term *merchant* is used to denote a person who receives the residual revenue after the agent receives his compensation. The term *trader* refers to both agents and merchants.)

Agents provided merchants with many trade-related services, including loading and unloading ships; paying customs, bribes, and transportation fees; storing the goods; transferring the goods to the market; and deciding when, how, and to whom to sell the goods and at what price and on what credit terms (Goitein 1967, p. 166). Agency relations enabled the Maghribi traders to reduce the cost of trade by better allocating risk by diversifying, by benefiting from agents' expertise, and by shifting trade activities across trade centers, goods, and time. Agency relations enabled merchants to operate as sedentary traders, thus saving the cost and risk of sea journeys. They also enabled traveling merchants to rely on agents to handle their affairs in their absence (Goitein 1967; Greif 1985, 1989).

The efficiency gain from operating through agents is impossible to assess quantitatively. But scholars have recognized the superiority of premodern trade systems in which cooperation through overseas agents prevailed over those in which it did not.[11] That the Maghribi traders themselves saw operating through agents as crucial for business success is reflected both in the extent to which they established agency relations and in their statements. "All profit occurring to me comes from your pocket," wrote one trader to his overseas agent. "People cannot operate without people," wrote another.[12]

3.2 The Commitment Problem and the Reputation-Based Community Enforcement Mechanism

[11] De Roover (1965, pp. 43-6, 70-5); Postan (1973, pp. 66-71); Lopez and Raymond (1955, p. 174)

[12] DK 22, b, line 18, Gil (1983a, pp. 97–106); TS 13 J 25, f. 18, Goitein (1967, p. 164). For the extent of agency relations through business associations, see Stillman (1970) and Michael (1965).

Transacting in agency services is characterized by a commitment problem. Letting an overseas agent conduct business with capital he does not own increases efficiency, but once the capital is in the agent's possession, the agent can embezzle it. Without a supporting institution, merchants, anticipating this opportunistic behavior, will refuse to operate through agents, and mutually beneficial exchanges in agency service will not be carried out. To surmount this commitment problem, there is a need for an institution enabling an agent to commit himself ex ante, before receiving the merchant's capital, to be honest ex post.[13]

The historical records implicitly indicate the existence of such an institution among the Maghribis, as agency relations were the rule rather than the exception. The *geniza* documents indicate that agency relations were characterized by trust. Despite the many opportunities for agents to cheat, only a handful of documents contain allegations of misconduct (Goitein 1973, p. 7).[14] How was the merchant-agent commitment problem resolved?

It was not resolved by using only family members as agents: in the sample used for this study, less than 12 percent of agency relations involved family members.[15] In some situations, a legal system can surmount a commitment problem inherent in a central transaction by linking it to a coercive (legal) transaction. Beliefs in legal sanctions deter misconduct. But the historical evidence suggests that this was not the case among the Maghribi traders. Many, if not most, of the agency relations reflected in the *geniza* were not based on legal contracts. Only a few documents indicate that commercial disputes between merchants and agents were brought before

---

[13] Were a merchant to sell the venture to an overseas agent, he would have to become the agent. Selling it to a local agent meant losing the advantages of an overseas agency.

[14] Misconduct is mentioned in less than 5 percent of the documents examined for this study.

[15] The evidence—commercial correspondence—is not likely to be biased against reflecting intrafamily overseas agency relationships. Information on specific merchants indicates that they hired nonrelatives. Naharay ben Nissim, a prominent merchant in Fustat, used more than 90 different agents. In the sample of contracts studied here, each merchant averaged 3.3 agents (excluding the two merchants with the largest number of agents—90 and 27—the average number of agents per merchant was 2.5 agents). This figure represents a lower bound, however,because many merchants probably had agents about whom no record was preserved. It is not possible to calculate the share of capital sent through agents who were family members. The evidence suggests, however, that such an evaluation would not alter the foregoing conclusion..

a court, and the operation of the court in these cases seems to have been expensive and time-consuming.[16]

The court also faced difficulties tracking down agents who had emigrated, and it was not structured to collect the information required to adjudicate disputes among traders regarding events that took place months before the trial, in faraway places (Greif 1989, 1993). Several months after the event, for example, a court could not verify the condition of goods upon their arrival, the price received for the goods, the amount of the bribe paid at the port, the cost of delivery, or whether the goods were stolen from the agent's warehouse. Moreover, Jewish law restricts the ability to sue agents. An agent entrusted to buy certain items cannot be sued for "bringing [to the merchant] an item worth 1 [dinar] for [which he charges the merchant] 100 [dinars]."[17] Indeed, in 1095 an agent who received 70 dinars reported having lost all but 20 dinars. The furious merchant, certain that he had been cheated, was unable to sue the agent, because his claim did not have any legal basis.[18]

The conviction of the furious merchant that the agent had cheated him was probably based on information that enabled him to monitor the agent only imperfectly. For diversification, traders were associated with many traders residing in different trade centers. It was customary for merchants to supply their business associates with trade-related information, which was crucial to business success.[19] Reciprocity probably prevented "free riding" on these information flows.[20]

---

[16] See discussion in Greif (1989). On the cost of litigation, see Bodl. MS Heb., a3, f 26, Goitein (1973, p. 97); TS 10 J 4, f. 4, Greif (1985, appendix, pp. 5–7); Bodl. MS. Heb., f. 42, Poznanski (1904, pp. 171–2); TS 20.152, Bodl. MS Heb., a3, f. 9, Gil (1983b, 2:724–32). Jews living in the Muslim world at the time could use the Jewish and the Muslim legal systems. The merchants' correspondence used here is equally likely to reflect the use of either system.

[17] Maimonides (1951, p. 208). See discussion in Greif (1989).

[18] TS 13 J 2, f. 5, Goitein (1967, p. 176).

[19] A Sicilian merchant, Jacob ben Isma`il, had at least five business associates in three different trade centers (see Greif 1985, p. 133). An important sedentary merchant like Naharay ben Nissim of Fustat had business relations with dozens of merchants from Spain to Syria (see Michael 1965 and letters to Naharay published in Gil 1983b, 3:96ff.).

[20] Trade-related information, including prices, ship arrivals and departures, and the general economic and political situation, appears in many *geniza* documents. See, for example, TS 20.76 and 13 J

Within the Maghribi traders group, these information flows, together with merchants' experience, reduced the asymmetry of information possessed by merchants and agents, enabling the merchants to monitor agents.[21]

These information flows also enabled agents to signal that they were honest. Just as modern firms hire auditors to establish the legitimacy of their financial statements, eleventh-century Maghribi agents generally conducted important business in the presence of other coalition members. In their reports they included the names of witnesses the merchant knew, thus enabling the merchant to verify the agent's report.[22]

The ability to monitor, however, was imperfect; a merchant could mistakenly conclude that an agent was dishonest. For example, around the middle of the century Maymun ben Khalpha of Palermo sent a letter to Naharay ben Nissim of Fustat. Discussing a conflict that Naharay had with one of his agents, Maymun makes clear that, in contrast to Naharay, he contends that the agent was honest and should not be accused of cheating.[23]

Ability to monitor agents is a necessary condition for surmounting the commitment problems by intertemporally linking merchant-agent transactions between a particular merchant and agent. Information regarding cheating is necessary for the merchant to adopt the strategy of hiring an agent each period as long as he is honest and never rehiring him if he ever cheats. Belief in this strategy can endogenously motivate the agent to be honest out of the desire to retain his position. To make the prospect of future employment sufficiently attractive to deter cheating, however, the merchant must create a gap between the expected lifetime utility the agent receives

---

15, f. 9, Goitein (1973, pp. 113–19, 320–22); TS 10 J 11, f. 22, a, lines 11–12. Cf. Goitein (1967, pp. 195, 201ff.) and additional references in Greif (1985, p. 95, n. 60). For the importance of information flow for commercial success, see Dropsie 389, a, lines 2–4, Gil (1983a, pp. 113–25); Michael (1965); Gil (1983b, 3:96ff.).

[21] For examples of such information, see DK 22, a, lines 11ff.., Gil (1983a, pp. 97–106); ULC Or. 1080 J 42, Gil (1983b, 3:300); TS Box Misc. 28, f. 225, Gil (1983b, 3:96–101).

[22]On the use of witnesses, see DK 13, sect. G; ULC Or. 1080 J 48; Bodl. MS Heb., a2, f. 17, all published in Goitein (1973, pp. 32, 92–93, 103). See also the discussion in Goitein (1967, pp. 168, 196) and Greif (1985, p. 143). In certain circumstances, Jewish law requires eyewitnesses. See Maimonides (1951, p. 214).

[23] See DK 22, b, line 5ff., Gil (1983a).

from working for the merchant and the agent's best alternative elsewhere. To do so, the merchant must provide the agent a per period premium; for example, he can pay him more than what he can earn elsewhere. Given these beliefs and premiums, a dishonest agent earns a short-run gain until he is caught, whereas an honest agent reaps a long-run gain by earning a premium each period.

This bilateral reputation mechanism relies on intertemporally linking merchant-agent transactions between the same parties.[24] Whenever the relationship between a merchant and an agent is expected to be terminated for exogenous reasons even though the agent was honest, more agency relationships in more situations can be entered into by linking transactions between different merchants and agents. Central to the associated reputation institution is an organization —a group of traders with a specific social identity ("coalition members")—who share information about agents' conduct. Members of this network share the beliefs that coalition merchants will employ only member agents and that each of them will reward his agent enough to keep him honest.[25] All coalition merchants, however, are expected never to employ an agent who cheated while operating on behalf of any coalition member.[26]

More agency relationships in a greater number of situations can be entered into in this case because, other things being equal, these beliefs reduce the premium that a merchant must pay an agent to keep him honest. These beliefs reduce the premium, because they lower the probability that a cheater will be able to earn the premium elsewhere. In addition, these beliefs enable merchants to employ agents for assignments that both parties know ahead of time will be of short duration. Because an agent who considers cheating a particular merchant risks his relations with all coalition members, the agent's lifetime expected utility is robust with respect to the length of his associations with a particular merchant. Hence the premium required to keep an

---

[24] See discussion of reputation mechanisms and references in Appendix C.

[25] The coalition, however, was not a monopsony in the usual sense of the term, since, as described later, a Maghribi trader usually operated simultaneously as a merchant and an agent.

[26] Chapter 9 extends the analysis to enable agents to be merchants. If an agent who is caught cheating operates as a merchant, coalition agents who cheat in their dealing with him are not expected to be considered by other coalition members to have cheated.

agent honest is unaffected by the expected duration of the agent's dealings with a particular merchant.

Theoretical considerations can generate many hypotheses; evidence is required to verify any postulate. The *geniza* contains direct evidence of the operation of a coalition. It suggests that a multilateral reputation mechanism governed agency relations; merchants conditioned future employment on past conduct, practiced community punishment, and ostracized agents who were considered cheaters until they compensated the injured party. The *geniza* further suggests that agents were ready to forgo current gains in order to sustain their good standing in the merchants' group.

Evidence of collective punishment within the coalition is found in two letters dated 1055. Abun ben Zedaka, an agent living in Jerusalem, was accused of (although not legally charged with) embezzling the money of a Maghribi trader. When word of the accusation reached other Maghribi traders, merchants as far away as Sicily canceled their agency relations with him.[27]

In the first decade of the eleventh century, Samhun ben Da'ud, a prominent trader from Tunisia, sent a long letter to his business associate, Joseph ben 'Awkal of Fustat. The letter reflects the traders' awareness of the importance of the prospect of future relationships as a motivating force. Joseph made this point clear when he made his future dealings with Samhun conditional upon his record: "If your handling of my business is correct, then I shall send you goods."[28] The letter reveals that future relations were conditioned upon past conduct—the essence of the reputation mechanism.

The same letter reveals the use of economic, rather than social, sanctions and the expectations for collective punishment among coalition members. Believing that Samhun had not remitted his revenues on time, Joseph imposed economic sanctions on him by not providing him with agent's services. He ignored Samhun's request to pay two of Samhun's creditors in Fustat and failed even to inform them of Samhun's request. By the time Samhun found out about it,

---

[27] TS 13 J 25, f. 12; TS 12.279. See also TS 8 J 19, f. 23. These letters are published in Gil (1983b, 3: 218–33).

[28] DK 13, a, line 41, Stillman (1970, pp. 267-75).

"their letters filled with condemnation had reached everyone." The letters caused Samhun to complain that "my reputation [or honor] is being ruined."[29]

The letter also reveals why agency relations were established and sheds light on their nature. It shows that economic interdependence, not internalized norms regarding mutual help or altruism, motivated the parties. Samhun cited two reasons for acting as Joseph's agent. The first was his desire to receive the agent's share of the profits. "You did not think that I should have a profit through you of even 10 dinars. Although you have made through me ten times as much." Elsewhere he mentions that he sold Joseph's pearls for 100 percent profit and added, "Should I not have taken one quarter of the profit?"[30]

The second reason why Samhun sought to maintain relations with Joseph was to increase the expected value of his capital. "What I do need is the benefit of your high position and for you to take care on my behalf," he writes. "It is my desire to avail myself of your high standing for those things which I send to you."[31] Note that the merchant is able to create a gap between the future utility stream of an honest agent and that of a cheater by controlling the expected income stream from the agent's capital. This correspondence thus suggests that agents received both a wage premium and a capital premium.

The deterrent effect of fearing the loss of one's reputation is clear from an incident described in a letter sent from Mazara, Sicily, in 1059. The writer had sold flax illegally (before the ships had arrived and the trading season officially opened) in Sfax, Tunisia, receiving an average price of thirteen dinars a load. By the time the ships arrived, the price had dropped to eight dinars a load and the buyers refused to pay the agreed-upon price. Eventually, the buyers paid, solely out of fear of harming their reputations. As the seller wrote, "We were lucky if not the [for their fear of losing their] honor we wouldn't have received a thing."[32]

---

[29] DK 13, a, lines 26 ff., Stillman (1970, pp. 267-75.).

[30] DK 13, b, lines 12–13, 20–21, Stillman (1970, pp. 267-75.).

[31] DK 13, a, lines 32 and 43, Stillman (1970, pp. 267-75.).

[32] Dropsie 389, b, lines 22ff., Gil (1983a, pp. 113–25). See also Bodl. MS Heb., a3, f. 26 and ULC Or. 1080 J 42, Goitein (1973, pp. 97, 92–5).

A letter sent around 1050 from Maymun ben Khalpha of Palermo to Naharay ben Nissim of Fustat also suggests that relations between a particular agent and merchant were of concern to other coalition members. Discussing a conflict that Naharay had with one of his agents in Palermo, Maymun writes, "You know that he is our [the Maghribi traders'] representative [so the conflict] bothers us all."[33] Another letter, sent around 1060, confirms the deterrent effect that fear of jeopardizing future relations had on opportunistic behavior. In this letter an agent justifies his actions, which caused some loss to the merchant, on the ground that he did not want people to say that he had contradicted the merchant's instructions.[34]

A letter sent in the middle of the eleventh century from a merchant in Palermo to Yeshu`a ben Isma`il in Alexandria further reveals the importance of reputation within the coalition.[35] The merchant describes how he handled the sale of two loads of pepper, one belonging to himself and the other belonging to his partner. The pepper price was very low: "I held [the pepper] until the time when the sailing of the ships approached in the hope [the price] would rise. However, the slump got worse. Then I was afraid that suspicion might arise against me and I sold your pepper to Spanish merchants for 133 [quarter dinars]. It was the night before the sailing of the ships pepper had became much in demand [because] boats [with buyers] arrived. Thus, [the pepper] was sold for 140–142. I took collateral for the sale of my pepper at 140–142. But brother, I would not like to take the profit for myself. Therefore, I transferred the entire sale to our partnership." The merchant decided to share the profits in order to maintain his reputation, even though he did not intend to do business with the partner in the future. "Settle my account with yourself and give the balance to my brother-in-law," he wrote, "for you are a very busy man." The merchant acted honorably solely to maintain his reputation with other coalition members.

The operation of a coalition is based on uncoordinated responses of merchants located in different trade centers. Hence it critically depends on a common cognitive system that ascribes meaning to various actions, particularly actions that constitute cheating. In other words, for the

---

[33] DK 22, b, line 5 ff., Gil (1983a, pp. 97–106).

[34] Bodl. MS Heb., d66, f. 60, a, margin, lines 7–9, Gil (1983b, 3:216).

[35] Bodl. MS Heb., a3, f. 13, Goitein (1973, p. 123).

threat of collective punishment to be credible, "cheating" must be defined in a manner that ensures collective response. If some merchants consider specific actions to constitute "cheating" while others hold a different opinion, the effectiveness of the collective threat is undermined.[36] The required coordination can be achieved by specifying an agent's obligations in an explicit contract, ideally a comprehensive contract. But given the state of communication technology and the uncertainty and complexity of trade during the eleventh century, detailed contracts entailed a high negotiation cost. If a merchant and an agent had had to agree on a contract before goods could be shipped to an agent, trade through agents would have been impractical.[37]

Indeed, the *geniza* reflects the extensive use of incomplete contracts, usually in the form of letters with instructions that involved no negotiation. "Do whatever your propitious judgment suggests to you," wrote Musa ben Ya`qub from Tyre, Lebanon, to his partner in Fustat some time in the second half of the eleventh century.[38] Merchants often authorized their agents to do whatever they deemed best if none of the prespecified contingencies occurred.

Incomplete contracts, however, undermine the operation of a coalition, because they do not define which actions represent cheating and allow agents to act strategically to take advantage of the incompleteness of the contract.[39] Theoretically, hierarchy (authority) relations may be used as a substitute for an ex ante comprehensive contract by assigning the merchant with the right to make all (ex post) decisions (O.Williamson 1985). Alternatively, culture may substitute for comprehensive contracts by specifying ex ante systematic rules of behavior.[40] Cultural rules can indicate what members of the organization should do after an unforeseen state

---

[36] For relevant theory, see Banks and Calvert 1989.

[37] The inappropriateness of comprehensive contracts in long-distance medieval trade is reflected in the difference between the Maliki and the Hanafi schools of law in Islam. See Udovitch (1970, pp. 208–09). For theoretical considerations of the inability to specify comprehensive contracts, see Hart and Moore (1999), Grossman and Hart (1986), O.Williamson (1985), Segal 1999.

[38] ULC Or. 1080 J 42, Goitein (1973, p. 94). For a similar situation in Europe, see Gras (1939, p. 80).

[39] Such behavior is reflected in Dropsie 389, Gil (1983a).

[40] See discussion in Camerer and Vespsalaninen (1987) and Landa (1988). Cf. Kreps (1990b).

of nature occurs. Hierarchy does not require ex ante learning of rules, but it does require ex post transmission of information between the parties; culture requires ex ante learning of the rules but no ex post communication.

Given the state of communication and transportation technology in the eleventh century, it is not surprising that the Maghribi traders did not rely on hierarchy.[41] Instead, they employed a set of cultural rules of behavior—merchants' law—that specified how agents needed to act to be considered honest in circumstances not mentioned in the merchant's instructions. The merchants' law was a commonly known rule among the Maghribi that served as a default contract between agents and merchants. Agents who were known not to have followed the merchants' law were considered cheaters.

The importance of the merchants' law in determining expectations about and attitudes toward an agent's behavior is reflected in the letter written by Maymun ben Khalpha to Naharay ben Nissim. In discussing the conflict between Naharay and his agent, Maymun justified the agent's actions by arguing that the agent "did something which is imposed by the trade and the communication [system; what you asked him to do] contradicts the merchants' law." (Another way to translate the term used is as "the way of the trade)." In another letter, a "very angry" merchant accused his business associate of taking "actions [that] are not those of a merchant."[42]

Little is known about the content of the merchants' law. The most convincing evidence for its existence and the process by which it emerged is found outside the *geniza*. In the middle of the twelfth century, Maimonides, an important Jewish spiritual leader who lived in Fustat, wrote in his legal code, "If [an agent] enters a partnership with another without specifying any terms, he should not deviate from the custom current in the land in regard to the merchandise they deal with" (Maimonides 1951, p. 223).[43] Similarly, the early medieval Islamic legal literature contains

---

[41] See DK 22, A, lines 9–11, Gil (1983a, pp. 97–106) for explicit statement indicating that it was impractical for an agent to await new instructions when an unspecified contingency occurred.

[42] DK 22, b, line 5ff., Gil (1983a, pp. 97–106); TS 12.434 l.7, Goitein (1967, p. 202, n. 50). See also Goitein (1967, p. 171).

[43] By that he did not mean a particular partnership form but partnership in general. This may indicate that the merchants' law was not specific to the Maghribi traders' coalition but was shared by a larger group. In the *geniza*, see DK 13, b. lines 7ff., Stillman (1970, p. 272); Dropsie 389, b, lines 22–23,

numerous instances in which systematic legal reasoning is suspended because of the "custom of the merchants" (Udovitch 1970, pp. 13, 250–9). Unfortunately, neither the legal literature nor the *geniza* documents reflect exactly how the merchants' law was formulated and changed.[44]

Within the Maghribi traders' coalition, the merchants' law promoted efficiency by providing a coordination device necessary for the functioning of the coalition, economizing on negotiating cost, and enabling flexibility in establishing agency relations. The merchants' law also imposed rigidity on the system, however, as the process of adjusting it was probably impeded by agents' concerns about what others would think about their actions rather than what the outcome of their actions would be. This is reflected in the words of Joseph ben Yeshua, an eleventh-century agent who wrote to a merchant that he could not act without written instructions because he did not wish that "people will say that I did something that I was not ordered."[45]

The historical record indicates the importance of a reputation-based institution. Informal, community-based contract enforcement mechanisms enabled the operation of a market in agency services. The historical evidence, however, raises many questions. Why was the community punishment self-enforcing? Why was the boycott of cheaters not undermined by agents' ability to seek employment with non-Maghribis? Why was the merchants' commitment to future employment of honest agents credible despite their ability to hire non-Maghribi agents?

3.3 The Model: The Agent Commitment Problem and the Multilateral Punishment Strategy

These questions raised in section 3.2 can be addressed using an explicit model. The analysis evaluates the claim that such coalition was possible and furthers our understanding of its operation. Constructing a model with which to examine the functioning of a contract enforcement institution in a specific historical episode presents a methodological problem.

Gil (1983a, pp. 113–25); TS 20.26, sect. I, Goitein (1973, p. 117).

[44]See, however, DK 22, on the a. margin, Gil (1983a, pp. 97–106); Goitein (1973, pp. 111–12); and Greif (1985, p. 136).

[45] Bodl. MS Heb., d66, f. 60, a. margin lines 7–9, Gil (1983b, 3:216).

Should the model's assumptions be restricted only to those reflected in the historical records? Or is any assumption about the model that does not conflict with the evidence legitimate? The position taken here (for reasons elaborated in Part IV) is that, to the extent possible, the model should be based on assumptions that are justifiable by the historical evidence, and it should account for the phenomena under consideration using the fewest additional assumptions.

The model presented here therefore does not impose the assumption that generates perhaps the most intuitive explanation for collective punishment—that merchants perceive an agent who cheated to be a "bad type," who will continue to cheat in the future.[46] There is no evidence to justify such an assumption, directly or indirectly, by indicating that an agent who had proved honest in the past was considered to be more likely to be honest in the future. On the contrary, there is evidence suggesting that merchants were likely to participate in collective punishment even when they believed that the agent was honest. In a letter quoted earlier, Maymun makes clear that he believes that Naharay's agent was honest and "should not be accused [of cheating]." Maymun feared that if the agent were openly accused, it would affect his relations with the agent, presumably since Maymun would have to participate in collective punishment.[47]

A model based on agents' types seems unable to provide a satisfactory explanation for some historical phenomena. For example, as discussed later, the Maghribis did not maintain agency relations with Jewish traders from Italy, although, ignoring agency cost, the Maghribis perceived such relations as very profitable. A model based on agents' types can account for this behavior, but it requires either making strategies contingent on social affiliations or assuming that members of one group could not verify whether a particular member of the other group ever cheated (implying that a non-Maghribi could not free-ride on the information generated by the Maghribis by observing their actions). Neither possibility is appealing. There is no reason to

---

[46] See discussion in Appendix C. Institutions with information sharing and collective punishment can also reflect such asymmetric information (see Kali 1999 and Annen 2003).

[47] DK 22, b, line 5ff., Gil (1983a, pp. 97–106). Similar considerations led to the rejection of a model in which costly participation in collective punishment is motived by punishing individuals who failed to punish a cheater (see Pearce 1995 and Kandori 1992).

believe that various Jewish trades "discriminated" against one another, and whether a particular individual was serving as an agent could easily be observed, because merchants could examine a ship's cargo, ownership, and destination (see Goitein 1967, pp. 336–7).

The model used here, which is based on the historical evidence, reveals another mechanism that can support collective punishment and account for other historical phenomenon. In this model, collective punishment is feasible due to the availability of information; it is self-enforcing due to the link between expectations about future hiring and the stream of rent required to keep an agent honest. To simplify the presentation, the model ignores the possible importance of imperfect monitoring.[48] The model is aimed at capturing the essence of intertransactional linkages and the associated institutional elements that generated behavior among the Maghribis.

Consider a perfect and complete information economy in which there are M merchants and A agents, each of whom lives an infinite number of periods. The Maghribi traders did not, of course, enjoy infinite life-spans, but relatives were considered morally responsible for one another's business dealings, and traders' sons became traders, serving as their parents' old-age insurance policies.[49] Hence the value of one's reputation did not diminish with old age. In considering the conditions under which reputation induces honesty, it therefore seems appropriate to assume an infinite horizon.

Assume (in accordance with the historical evidence) that there are less merchants than agents, $M < A$.[50] Agents have a time discount factor δ, and an unemployed agent receives a per

---

[48] To capture the asymmetry and imperfectness of information, as well as commercial uncertainty, the model presented here can be extended as follows: revenue is observed only by the agent and is a random variable x with domain $[a,b]$. The agent reports a revenue realization $y \in [a, b]$. A wage is a contract that is a function of the agent's report, $w{:}[a, b] \to [a, b]$, $w(y) \leq y \; \forall \; y$. The merchant observes the actual realization with probability $f(y, x)$, where $1 > f(.) > 0$, $\forall \; y \neq x$ (information asymmetry), and $f(.) > 0$ when $x = y$ (imperfect monitoring). The path of play in imperfect monitoring models is characterized by episodes of noncooperation; in some periods the agent is punished for perceived cheating by not being hired (see Appendix C, section C.2.7).

[49] See Goitein (1973, p. 60). Goitein (1978, pp. 33ff.) notes that "both the government and public opinion were prone to hold a father, or brother, or even more distant relative responsible for a man's commitments, although strict law, both Islamic and Judaic, did not recognize such a claim."

[50] More specifically, the historical evidence indicates that merchants were not deterred from terminating agency relationships fearing to be unable to hire an alternative agent.

period reservation utility of $\bar{w} \geq 0$. In each period, an agent can be hired by only one merchant and a merchant can employ only one agent. Matching is random, but a merchant can restrict the matching to a subset of the unemployed agents made up of agents who, according to the information available to the merchant, have previously taken particular sequences of actions.[51]

A merchant who does not hire an agent receives a payoff of $\kappa > 0$. The gross gain from cooperation is γ. A merchant who hires an agent decides what wage ($W \geq 0$) to offer the agent. Because an employed agent held the merchant's capital, it is appropriate to assume that an agent is ensured of receiving his wage. An employed agent can decide whether to be honest or to cheat, and his actions are public information. If the agent is honest, the merchant's payoff is $\gamma - W$, and the agent's payoff is W. If he cheats, his payoff is $\alpha > 0$ and the merchant's payoff is $\gamma - \alpha$. It is assumed that $\gamma > \kappa + \bar{w}$ (cooperation is efficient); $\gamma > \alpha > \bar{w}$ (cheating entails a loss, and an agent prefers cheating over receiving his reservation utility); and $\kappa > \gamma - \alpha$ (a merchant prefers not to hire an agent and to receive κ over being cheated or paying a wage as high as the amount the agent can cheat him by).

After the allocation of the payoffs, each merchant can decide whether to terminate his relations with his agent. There is a probability τ, however, that a merchant is forced to terminate agency relations. The need to shift commercial operations over places and goods—and the high level of uncertainty of commerce and life during the eleventh century—curtailed a merchant's ability to commit himself to future wages or employment. Hence the model assumes a stationary wage scheme (which was indeed practiced among the Maghribis) and a limited ability to commit to future employment.[52] Finally, wages were neither politically nor legally determined, and there is no evidence of collusion in wage determination. Accordingly, the analysis assumes that no subgroup is organized in a way that affects wage determination.

---

[51] The following assumes that the probability of rematching with the same agent equals zero for all practical consideration.

[52] For an efficiency wage model in which this result is derived endogenously, see MacLeod and Malcomson (1989). Their approach could be utilized here as well; it is omitted to maintain simplicity. Levin (2003) has established that self-enforcing contracts in repeated settings with moral hazard will generally be simple and stationary.

Analyzing this model – a version of a one-sided prisoner's dilemma game – highlights why the collective punishment was self-enforcing among the Maghribis. Consider a multilateral (collective) punishment strategy combination according to which a merchant offers an agent a wage W^*, rehires the same agent if he was honest (unless forced separation occurred), fires the agent if he cheated, never hires an agent who has ever cheated any merchant, and (randomly) chooses an agent from among the unemployed agents who never cheated if forced separation occurred. An agent's strategy calls for being honest if paid W^* and for cheating if paid less than W^*. Is a multilateral punishment strategy a subgame perfect equilibrium? Will a merchant punish an agent who did not cheat him?

To address these questions it is necessary to consider the factors determining the wage, W^*, that will be offered by the merchants. Let h_h denote the probability that an unemployed honest agent (i.e., an agent who was honest when last employed) will be rehired. Let h_c denote the probability that an unemployed cheater (i.e., an agent who cheated when last employed) will be rehired. Proposition 3.1 presents the relationship between the model's parameters, these probabilities, and the lowest wage for which an agent's best response is to be honest.[53]

Proposition 3.1: Assume that $\delta \in (0, 1)$, $h_c < 1$. The *optimal wage*, the minimum (symmetric) wage for which, if offered by all merchants, an agent's best response is to be honest, is $W^* = w(\delta, h_h, h_c, \tau, \bar{w}, \alpha) > \bar{w}$, where w is monotonically decreasing in δ and h_h and monotonically increasing in h_c, τ, $\bar{w}$, and α.[54] (Proof is given in annex 3.1.)

Under a multilateral punishment strategy, what motivates an agent to be honest is the carrot of receiving a premium over his reservation utility and the stick of being fired. If the difference between the present value of the lifetime expected utility of an unemployed cheater and an employed agent is higher than the one-period gain from cheating, an agent's best response is to be honest. The optimal wage thus decreases if an honest agent is more likely to receive future wage premiums (higher h_h), can gain less by cheating (lower α), is more likely to remain

[53] This specification enables the optimal wage under both the multilateral and bilateral punishment strategies discussed later to be examined.

[54] More precisely, this monotonicity is weak at some neighborhoods of the extreme values of the parameters.

employed if he is honest (lower τ), has worse opportunities elsewhere (lower $\bar{w}$), and has a smaller chance of being hired if he cheats (lower h_c). Further, the optimal wage decreases as an agent values future income more (higher δ), because rewarding for honesty and punishing for cheating is done in the future.

For a multilateral punishment strategy to constitute a symmetric subgame perfect equilibrium, each merchant should find it optimal to hire agents. On the equilibrium path, this condition means that the wage is set low enough—that is, $W^* = w(., h_c, h_h) \leq \gamma - \kappa$, where $h_c = 0$, and $h_h = \tau M/(A - (1 - \tau)M)$.[55] Assume that this condition holds. Will a merchant find it optimal to punish an agent who did not cheat him? When switching agents does not impose any cost—as assumed here—merchants may as well punish a cheater; hence the multilateral punishment strategy is a subgame perfect equilibrium. Having the credibility of multilateral punishment rest on a knife-edge result, however, is unsatisfactory. Clearly, Maymun be Khalpha considered that punishing the Sicilian agent was costly. Therefore, a more relevant question is whether the multilateral punishment strategy motivates a merchant *strictly* to prefer hiring an honest agent over a cheater.

As proposition 3.2 demonstrates, a merchant strictly prefers to hire an honest agent under the multilateral punishment strategy merely because a cheater is not expected to be hired by other merchants. An honest agent is expected to be hired in the future, but an agent who has ever cheated is not. Because the optimal wage decreases in the probability of future hiring, a cheater's optimal wage is higher than an honest agent's wage. Hence each merchant strictly prefers to hire an honest agent. The uncoordinated response of all the merchants and the interrelations between their expected future behavior and an agent's optimal wage as perceived by an individual merchant ensure solidarity of incentives. The possibility of forced separation links the optimal wage a particular merchant has to pay his agent and the agent's expected future relations with other merchants. This link increases the optimal cheater's wage above an honest agent's wage, because punishments are independent of the agent's past conduct while rewards are not. Hence merchants find it optimal to follow the multilateral punishment, *despite* the fact that the agent's

[55] These probabilities are induced by the strategies.

strategy does not call for cheating any merchant who violated the collective punishment, and *despite* the fact that cheating in the past does not indicate that the agent is a "lemon." Hence it is reasonable that Maymun was concerned about Naharay's interpretation of his agent's actions, because open accusation would have initiated an uncoordinated response that would have affected Maymun's business with that agent.[56]

Proposition 3.2: Assume that $\delta \in (0, 1)$ and $h_c < 1$. Under a multilateral punishment strategy, a merchant *strictly* prefers to hire an honest agent. (Proof in annex 3.1.)

3.4 The Maghribi Traders' Coalition: Theory and Indirect Evidence

The historical anecdotes presented in section 3.3 suggest that contract enforcement among the Maghribis was achieved by collective punishment. The model and its equilibrium analysis lend support to the conjecture about the importance of this reputation mechanism by indicating that it is logically consistent. But we can do better than that. It is possible to substantiate the hypothesis further by considering the model's implications. A coherent explanation of historical observations can be advanced based on the assertion that a coalition governed agency relations. Predictions based on this assertion can be generated and confirmed by historical records.

The record is rich in facts that should be explained. The Maghribis were the descendants of merchants who lived in the `Abbasid Caliphate, centered in Baghdad until the first half of the tenth century. Military conflicts and political instability caused these merchants to emigrate, during the tenth century, mainly to Tunisia, which prospered at the time, under the control of the Fatimid Caliphate. With time these traders extended their trade from Spain to Constantinople. While the agency relations required for this expansion could have been established with non-Maghribi traders (Jewish or Muslim), evidence of such relations is rare. Instead, members of the Maghribi traders' group emigrated, to Spain, Sicily, Egypt, and Palestine. For generations

[56] One document (Bdl. MS Heb., a2, f. 17, sect. D, Goitein 1973, p. 104) also reveals another, related way through which the expectations for collective punishment rendered it self-enforcing. Because traders usually acted as both merchants and agents, they maintained "open accounts" with other traders - that is, accounts that were cleared only periodically. When an agent was rumored to be in trouble, traders feared that he would not be able to pay his debts. As a preventive measure, they ceased sending him goods and held on to the money they owed him.

members of these colonies maintained agency relations with the descendants of other Maghribi traders.[57]

Because the Maghribis lived in the Arab world for centuries, they adopted its customs and language. Hence, emigration outside the Arab sphere of influence was culturally and materially difficult. Indeed, the Maghribis did not emigrate to the emerging trade centers of Italy, despite their perception that trade with the Christian world was very profitable.[58] This perception is reflected, for example, in the words of a merchant from Palermo, who complained around 1035 that *even* the Rums (in this case, Christians from the Latin world) were not ready to buy the inferior black ginger.[59] Arguably, it was often easy to sell them inferior goods for prices noone else was willing to pay. Despite the perceived profitability of this trade, Maghribi traders did not establish agency relations with Jewish traders from Italy who were active during this period. The communities within which the Maghribi traders operated maintained communal ties with the Italian Jewish communities, and no political restrictions hindered cooperation between the Maghribis and the Italian Jews. Yet there is no evidence of agency relations between the Maghribis and Jewish traders from the Christian world.

In the trade centers to which the Maghribi traders emigrated, a well-established Jewish community already existed, into which the Maghribi traders integrated. But as long as they were active in long-distance trade, they preserved a separate social identity. This identity is reflected in documents in which they refer to themselves as "our people, the Maghribis, the travelers (traders)."[60]

---

[57] Goitein (1967, pp. 156–9, 186–92); Gil (1983b, 1:200ff.); Greif (1985, pp. 124–7).

[58] See, for example ,TS 8 Ja I, f. 5, Goitein (1973, pp. 44–5); Goitein (1973, p. 211); and Greif (1989).

[59] Dropsie 389, b, lines 6 ff., Goitein (1973, p. 45). See also Bodl. MS Heb., c28, f. 11, lines 11–13.

[60] Gil (1971, pp. 12–15; 1983b, 1:215, 223); Goitein (1967, pp. 30–4, 148–9, 157); Greif (1985, p. 153, n. 32); and see, for example, DK 13, sect. G, F, Goitein (1973, p. 32); TS Box Misc. 25, f. 106, a, line 9, Gil (1983b, 2:734); TS 13 J 26, f. 24, b, lines 3–5; TS Box Misc 25, f. 106, line 9, Gil (1983b, 2:601, 734).

The Maghribis operated in the Mediterranean sea during the eleventh century, until Italian and more generally European naval, military, and commercial ascendency in that area drove merchants from the Muslim world away. The Maghribis then turned to the Indian Ocean trade, until the end of the twelfth century, when they were forced by the Muslim rulers of Egypt to withdraw.[61] At that point they integrated with the larger Jewish communities and vanished from the stage of history.

These historical observations raise intriguing questions. Why did the Maghribis not establish seemingly profitable agency relations with non-Maghribis? How can the governance of agency relations by a coalition and the possibility of establishing an agency with nonmembers be reconciled? The possibility of hiring nonmember agents seems to undermine the member merchants' commitment to hire honest member agents in the future, and it seems to undermine the effectiveness of collective punishment, because agents can potentially forge agency relations with nonmember merchants. What ensured the closeness of the coalition? Why was it self-enforcing and hence sustainable?

To support the hypothesis that a coalition governed agency relations, these issues should be explained in a way that is consistent with the assertion that agency relations were governed by a coalition. Furthermore, theoretical insights consistent with this assertion should be able to tie the Maghribis' immigration to Tunisia with the emergence of the coalition, as well as account for the fact that the Maghribis retained their social identity only as long as they were active in long-distance trade.

To address these questions, we need to examine the relations between coalition and efficiency. A coalition increases efficiency relative to a situation in which agency relations are governed by the bilateral punishment strategy. This bilateral strategy is identical to the multilateral punishment strategy, except that merchants do not condition their hiring on past conduct (because they do not have information regarding past actions, they do not expect others to make hiring conditional on that information, or they do not observe the wage paid to the agent and believe that cheating reflects underpayment).

---

[61] For the Maghribis' trade in the Indian Ocean, see Fischel (1958).

Under a bilateral punishment strategy, merchants do not hire agents in situations in which they would hire agents under a multilateral punishment strategy. Consider, for example, the case in which each merchant can commit himself to hire an agent for only one period ($\tau = 1$). Under this condition, in a bilateral punishment strategy, for any finite wage agents will cheat. Agents are thus never hired. In contrast, under a multilateral punishment strategy, an agent takes into account the implications of cheating a particular merchant on his future employment with other merchants. The optimal wage will be finite and may be low enough to support cooperation.

Indeed, agency relations among the Maghribis were extremely flexible, as merchants operated through several agents at the same time and even at the same trade center. Agency relations were initiated and canceled with ease, depending on merchants' needs (Stillman 1970; Greif 1985).

Proposition 3.3 indicates that in general a multilateral punishment strategy supports cooperation when a bilateral punishment strategy fails to do so, due to the limited ability of each merchant to commit himself to rehire an honest agent by decreasing the probability that a cheater will be rehired, h_c.

Proposition 3.3: For ease of presentation, suppose that the agents' time discount factor approaches one ($\delta \to 1$). Define a as the ratio of agents to merchants, $a = A/M$. Recall that $\bar{w} < \alpha$ and $a > 1$. Given a, cooperation is feasible for all $\tau \in [0, 1]$ if and only if $\gamma - \kappa \geq (a - 1)\bar{w} + \alpha + \epsilon$, $\forall\ \epsilon > 0$ under a bilateral punishment strategy but if and only if $\gamma - \kappa \geq a\bar{w} + \epsilon$, $\forall\ \epsilon > 0$ under a multilateral punishment strategy. Given τ, cooperation is feasible for all $a \geq 1$ if and only if $(\gamma - \kappa) \geq \alpha + \epsilon$, $\forall\ \epsilon > 0$ under a bilateral punishment strategy but if and only if $(\gamma - \kappa) \geq \bar{w} + \epsilon$, $\forall\ \epsilon > 0$ under a multilateral punishment strategy. (Proof is given in annex 3.1.)

A multilateral punishment strategy enhances efficiency, because it enables cooperation when the ability of each merchant to commit to hire an agent in the future is limited. As long as the ability of a merchant to commit to hiring an agent in the future is less than perfect, a coalition decreases the wage, W^*, relative to the wage that prevails when a bilateral punishment strategy governs agency relations. This reduction reflects a decrease in the probability that a cheater will be hired, (h_c), and an increase in the probability that an honest agent will be hired (h_h). This wage reduction enhances efficiency by making agency relations profitable in situations in which the

total gain from cooperation is relatively low (γ is small). Although in such cases cooperation is efficient, it will be initiated only if it is profitable to a merchant, that is, $W^* \leq \gamma - \kappa$. Because the wage under a multilateral punishment strategy is lower than under a bilateral punishment strategy, more cooperation will be initiated. The wage reduction and the enhanced efficiency imply that organizing agency relations in a coalition increases member merchants' profits and may increase the lifetime expected utility of an honest agent who is a coalition member relative to that of an honest agent under a bilateral punishment strategy.

Efficiency gains generated by a coalition encourage its emergence; the coalition rewards member merchants and agents in a manner that encourages agency relations among coalition members. Hence by affecting efficiency and profitability, the beliefs that members will hire and be hired only internally can be self-enforcing: member merchants are motivated to establish agency relations with member agents, and member agents are better off being employed by member merchants.

Additional factors also contribute to this result. Expectations with respect to future hiring, the benefits provided by the network for information transmission, and strategic considerations discouraged members from initiating agency relations with nonmembers, and they discouraged nonmembers from initiating agency relations with members.

To see the impact of these factors, consider an economy with two identical coalitions. By definition, coalition members are not expected to establish intercoalition agency relations. Will these expectations be self-enforcing? A merchant will initiate intercoalition agency relations only if it is expected—that is, the institutionalized beliefs hold—that the other coalition's merchants will employ a multilateral punishment strategy against a member agent who cheated a nonmember merchant. Otherwise, the merchant strictly prefers to establish intracoalition agency relations because the optimal wage in intercoalition agency relations is $w(., h_c = h_h > 0)$, which, by proposition 3.1, is strictly higher than the optimal wage in intracoalition agency relations, $w(., h_c = 0, h_h > 0)$. For this wage differential to exist, it is sufficient that the merchant be uncertain about whether a multilateral punishment strategy will be applied in intercoalition relations.[62]

[62] The formal analysis of this argument is presented in propositions 9.4 and 9.5, which consider a more general issue the implications of different beliefs regarding behavior off-the-path-of-play on the

A merchant is likely to be uncertain about whether a multilateral punishment strategy will be applied in intercoalition relations due to information barriers between coalitions and strategic considerations. The fact that within a coalition each trader is known to others enables informal information flows, which the agent does not control, to facilitate monitoring and to inform traders about cheating. This mechanism does not function in intercoalition agency relations. Furthermore, coalition members are strategically motivated to ignore an outsider's accusations concerning the conduct of a coalition member agent. If the coalition members simply take the word of an outsider, an agent is vulnerable to blackmail by nonmembers, which reduces his lifetime expected utility as an honest agent. This reduction comes at the expense of member merchants, because it increases the optimal wage. Hence coalition members find it optimal to ignore outsiders' accusations. In contrast, insiders' accusations are not likely to be ignored, because they can be assessed more accurately and an insider merchant puts his own reputation on the line in accusing an agent. Khalluf ben Musa seems to have regretted ignoring insiders' accusation when he wrote to his partner in response to the accusation that he had retained revenues received for the partner's goods "had I listened to what people say, I never would have entered into a partnership with you."[63]

As a multilateral punishment strategy does not apply in intercoalition relations, the wage required to keep an agent honest in intercoalition agency relations is higher than the intracoalition wage. Merchants are thus discouraged from establishing intercoalition agency relations, and the expectations that intercoalition agency relations will not be initiated are self-enforcing. Note that this result holds even in some situations in which these intercoalition relations are more efficient. More precisely, intercoalition agency relations will not be established if the increase in the gains from cooperation does not compensate a merchant for the wage increase.

Expectations with respect to future hiring, the nature of the networks for information transmission, and strategic considerations ensure the self-enforceability of a coalition. These

---

motivation to establish intereconomy agency relations.

[63]Bodl. MS Heb., a3, f. 13, sect. B, Goitein (1973, p. 121). See also DK 13, sect. G; ULC Or. 1080 J 48; Bodl MS Heb. A2, f 17, Goitein (1973: 32, 92-93, 103). Goitein (1967a: pp. 168, 196); Greif (1985: 143).

factors encourage member merchants to hire only member agents and discourage member merchants from hiring nonmember agents. They enable member merchants to commit to hire only member agents, even if efficient agency relations can be established with nonmembers. At the same time, these factors make collective punishment effective, by discouraging nonmember merchants from hiring member agents, thus enabling member agents to commit themselves not to enter agency relations outside the coalition. By discouraging intercoalition agency relations, these factors contribute to making the beliefs on which the coalition rests self-enforcing. Hence once a coalition is formed through some historical process, agency relations will be established only among the traders about whom expectations were initially crystallized.

These theoretical observations suggest that the informal social networks for information transmission, which became available to the Maghribis in the process of immigrating to Tunisia, enabled them to support agency relations based on a multilateral punishment strategy.[64] This immigration process determined the social identity (position) of the individuals with respect to whom expectations of collective punishment and future hiring were established. Within the resulting coalition, information regarding the circumstances an agent faced was essentially free, because it was obtained as a by-product of commercial activity and passed on along with other commercial correspondence. The fact that the marginal cost of obtaining this information was essentially zero is important, because it made credible the merchant's claim that he would monitor his agents. Without such monitoring, the reputation mechanism could not have functioned.[65]

Once these beliefs were institutionalized—once the Maghribi traders' coalition was formed—only descendants of Maghribis were perceived by others to be members, and hence only they could become members. The factors that encouraged intracoalition agency relations and discouraged agency relations with nonmembers made membership a valuable asset. For this

---

[64] It was a necessary but not sufficient condition. See Chapter 9 regarding the importance of cultural beliefs in leading to the coalition.

[65] According to the theory advanced here, agents do not cheat. Thus, if monitoring agents is costly, it is not credible. Knowing that the merchants will not monitor, agents will cheat. Anticipating this, a merchant will not employ agents to begin with.

reason, the descendants of Maghribi traders continued to be active in long-distance commerce as members of the Maghribi traders' coalition. This, in turn, implied that each trader had a horizon long enough to render a reputation mechanism effective, because his children could have been punished if he defaulted.

As the Maghribis expanded the geographical scope of their trade, the profitability of intracoalition agency relations became high enough to encourage emigration and the establishment of colonies in other trade centers. Because Maghribi merchants were motivated to employ other coalition members, they were able to commit themselves to future employment of Maghribi agents. This ensured the emigrants that they would be compensated for the cost of emigration. Emigration to Italy was more difficult culturally and therefore forgone. Nonmember Italian Jews were not employed as agents, despite the common religion and the potential gains from trade with Italy, since the additional gains from establishing agency relations outside the coalition did not compensate for the relatively high agency cost.[66]

The Maghribi traders' social structure was an organization that provided them with the initial information-transmission mechanism required for the emergence of an economic institution—the Maghribi traders' coalition. This economic institution for the governing of agency relations provided the interactions required to sustain the social structure, while the Maghribis' social identity provided the means to coordinate expectations required for the functioning of the coalition. When the Maghribis were forced by the Muslim rulers to cease operating in long-distance trade and their coalition ceased to function, the motivation for social interactions diminished, their social structure lost its vitality, and the Maghribi traders assimilated into the broader Jewish community.

---

[66] Goitein (1967) conjectures that the lack of evidence of such relationships reflects selection bias, as the correspondents did not pass through Egypt. Yet the *geniza* is rich in documents reflecting agency relationships with agents in North Africa, Sicily, and Spain. In many cases, we learn about trade in Spain from communication with agents in Sicily and this island was also on route to Italy. Consistent with the argument advanced here, the Maghribis did integrate with other communities in matters unrelated to agency relationship. Such non-agency-related relationships are well reflected in the *geniza*. (Goitein 1967.) Yet agency relationships rarely appear. For example, in the letters of Naharay ben Nissim, the most important Maghribi trader in Fustat in the mid-eleventh century, only two of the ninety-seven different traders mentioned were Muslims (Michael 1965).

The discussion of the credibility of the collective punishment and endogenous information flows suggests that as long at the Maghribi traders' coalition survived, its functioning crucially depended on maintaining an appropriate size. The credibility of collective punishment rests on the coalition being sufficiently large that one agent can be used as a substitute for another if the latter cheats. But, everything else being equal, a larger coalition implies a slower circulation of information and hence delayed punishment (which can be captured in the model by making the time discount factor an increasing function of the coalition size).

The conjecture about the operation of a multilateral reputation mechanism gains further support from illuminating the rationale behind patterns of bookkeeping and employment of agents among the Maghribi traders. Agency relations among the Maghribis resembled the relations between a modern firm and its workers in that typically no explicit legal commitment governed the length of the relationship. When a commitment was made, it was for a short period of time. The duration of agency relations varied from a single season to several generations, with sons replacing their fathers as agents.[67] The Maghribi traders used a per trade venture rather than multiventure accounting system, in which the income and expenses associated with each trade venture were detailed (Goitein 1967, pp. 178, 204 ff.).

These trade practices are consistent with the operation of a reputation mechanism within a coalition. Whenever a reputation mechanism is employed, a merchant may prefer short-term contracts, because the shorter the contract, the sooner the merchant can discover deviation and the less he will have to pay to keep an agent honest. A per venture accounting system is more efficient than a multiventure accounting system whenever a reputation mechanism is employed, because it facilitates comparing agents' reports with any relevant information.

The historical evidence presents other puzzling questions. Why did most Maghribi traders operate as both merchants and agents? Why did the Maghribis employ agents through particular contractual forms? Why did different practices prevail among other comparable traders' groups,

---

[67] Goitein (1967, pp. 169–70, 178) and Greif (1985, p. 133). In the Italian trade cities, *commenda* relations were also of short duration (see, e.g., Lopez 1952, p. 323). Sons did not inherent their fathers' businesses per-se but members of the younger generation began providing agency services to each other.

such as the Italians, who also operated around then? Were efficiency and profitability sufficient to lead to the emergence of the coalition? These questions are better addressed in the context of a comparative study between these groups, presented in Chapter 9.[68]

3.5 Concluding Comments

The Maghribi traders' coalition mitigated problems of contract enforceability and coordination that arose in complex trade characterized by asymmetric information, slow communication technology, inability to specify comprehensive contracts, and limited legal contract enforceability. Within the coalition, information flows enabled monitoring and made cheating known, while a merchants' law coordinated responses. The multilateral punishment, the value of the information flows for commercial success, and the importance of the merchants' law as a substitute for comprehensive contracts generated wage and capital premiums. Receiving these premiums was conditional on past conduct, while intergenerational transfers ensured a horizon long enough to support the operation of a reputation mechanism. Because the present value of the premiums was larger than what an agent could gain by cheating, agents could credibly commit themselves to be honest, and merchants could trust them.

The coalition reflects intertransactional linkages and institutional elements exogenous to each individual whose behavior it influenced. It illustrates the importance of going beyond the confines of classical contract theory in economics (surveyed in Hart and Holmstrom 1987), which explores how bilateral contracts mitigate contractual problems. The analysis of the Maghribis highlights the importance of the social context in mitigating bilateral contractual problems. The Maghribi traders' group was a means of linking information-sharing transactions among merchants and fostering the personal familiarity required to render a threat of collective punishment credible. Beliefs about future hiring and collective punishment among the members of this organization linked each merchant-agent transaction with agents' future transactions with all Maghribi traders. The Maghribi traders' group and the commonly known beliefs in collective

[68] The historical evidence presents another puzzling question: why did Maghribi agents not cheat and begin to operate as merchants? The model presented here ignores this issue by assuming that an agent consumes whatever he appropriates. I address this issue in Chapter 9.

punishment were exogenous to each individual trader, while each member's best response to them was to maintain his affiliation with the group and to participate in collective punishment. Hence the coalition was self-enforcing and the threat of collective punishment credible. The commonly known merchants' law lent credibility to the threat of collective punishment by providing the unified interpretation of actions and facilitating a coordinated response.

The institution that supported trust among the Maghribi traders was composed of several inter-related social factors, rules, beliefs, and organizations. Togther, these institutional elements enabled, guided and motivated a particular regularity of behavior, namely, the intra-group hiring of agents and honesty. Rules provided the common cognition, coordination, and information that enabled and guided behavior in the related transactions. They enabled traders to make informed decisions by providing micro-foundations of behavior. Rules specified, for example, the structure of the situation, who held membership in the Maghribi traders' coalition, how one gained relevant information, which actions constituted cheating, how one complained about it, and what behavior was expected of merchants and agents if cheating transpired.

Beliefs motivated the traders to follow the behavioral instructions provided by these rules. It was common knowledge that the prevailing internalized and behavioral beliefs were that merchants would hire only Marghribi agents and would participate in collective punishment, and that agents would not be employed by non-members and would be honest. Given this, it was optimal for most traders, in general, to follow the behavioral instructions provided by the rules. The Maghribi traders group constituted a means for rules to be specified and information generated and disseminated. In particular, information flows within the group and the common interpretation of agents' actions increased the set of situations in which agents could credibly commit to be honest.

The origin and size of the coalition do not reflect the function it served. Rather, they reflect the relationships between an immigration process, the resulting social group, and, as elaborated in Chapter 9, historically inherited cultural beliefs. The social identity of and network for information transmission within an immigrant group determined the coalition's initial size. In the resulting coalition, the original social identity served as a signal that coordinated actions and expectations. By promoting agency relations and information transmission among a particular

group of individuals, the economic institution that governed agency relations preserved the initial social structure, which in turn determined the boundaries of the economic institution.

By reducing agency costs and other transactions costs, the coalition promoted efficiency and profitability among its members. It provided the foundations for the operation of a market in agents' services, enabling merchants to operate through agents even when the cost of establishing agency relations between a particular merchant and an agent in isolation would have been prohibitively high. The merchants' law economized on negotiation cost, governed the transmission of information and the provision of services, and substituted for comprehensive contracts between particular agents and merchants.

Despite these advantages, the coalition seems not to have been an optimal institution in the sense of enabling all gains from agency relationships given the period's enforcement, communication, and production technology. Specifically, the coalition was not dynamically efficient. The same factors that ensured its self-enforceability prevented it from expanding in response to welfare-enhancing opportunities.[69] The merchants' law potentially introduced another distortion, as its modification seems to have been done in a manner that did not ensure optimal changes. Within a coalition, agents are more concerned about the interpretations of their actions by other members than about the outcomes of their actions. Hence their actions, while aiming to maximize their expected utility, do not necessarily maximize total profit. The introduction of some form of leadership might have mitigated these distortions, albeit possibly at the cost of introducing others.

The analysis of the Maghribi traders' coalition illustrates the importance of contract enforcement institutions in the operation of markets. This nonmarket institution provided the foundation for a market in agency services, thereby contributing to the integration of interregional product markets. The nature of nonmarket institutions affects the cost, and possibly the feasibility, of trade, thereby affecting the ability to exchange and the process of market integration. As market integration is commonly believed to be a key to economic growth,

[69] In general, networks with collective punishment can be inefficient, particularly when entailing negative externalities on nonmember merchants. The magnitude of the effect depends on how the contractual problem that the network mitigates would have been resolved in its absence. For relevant theoretical analyses, see, for example, Kali (1999) and Dasgupta (2000).

institutional analysis of nonmarket institutions, their relationships to social and business networks, and their relations to market integration is a key to advancing our understanding of the processes of economic growth.

Social networks and ethnic groups play an important role in facilitating contract enforcement in the absence of the law, as many sociologists, anthropologists, and economists have noted (Macaulay 1963; Furnivall 1956; Landa 1978; Granovetter 1985; Homans 1961; Nee and Ingram 1998). These analyses, however, tend to take the social network and the credibility of punishment within it as a given. The analysis of the Maghribis highlights that fully understanding the nature and implications of economic institutions related to these networks and groups requires understanding the dynamic interplay between the social structure and the related economic institution. Understanding this interplay in various historical periods and economies is likely to provide an important supplement to the study of the institutional foundations of markets provided by the state and the interrelationships between the two.[70]

[70] For important recent contributions and surveys of such findings, see Rauch (2001); Casella and Rauch (2002); McMillan and Woodruff (1999); Kranton and Minehart (2001); and Chwe (2001). See Greif (1994b) and McMillan (2002) regarding the relationships between private and public contract enforcement institutions.

Annex 3.1

Proof of Proposition 3.1

For a given h_c and h_h, to show that playing honest is optimal for the agent, it is sufficient to show that he cannot gain from playing cheat one period if offered W^*. Accordingly, let V_h denote the present value of the lifetime expected utility of an employed agent who, whenever hired, plays honest, V_h^u denote the lifetime expected utility of an unemployed honest agent, and V_c^u the lifetime expected utility of an unemployed cheater (who will play honest in the future if hired). Note that the last two expressions take into account only income from the next period and on which is the first period of unemployment. These lifetime expected utilities are:

$$V_h = W^* + \delta(1 - \tau)V_h + \tau V_h^u,\ V_i^u = \delta h_i V_h + \delta(1 - h_i)(\bar{w} + \delta V_i^u),\ \ i = h, c.$$

Cheating once yields $\alpha + V_c^u$. An agent will thus not cheat if $V_h \geq \alpha + V_c^u$. Substituting and rearranging terms show that an agent's best response is playing honest if and only if $W \geq (T - \delta\tau H_h)[\alpha/(1 - \delta H_c) + \delta\bar{w}(P_c/(1 - \delta H_c) - \tau P_h)] = W^*$, where $T = 1 - \delta(1 - \tau)$; $H_i = h_i/(1 - \delta^2(1 - h_i))$, $i = h, c$; $P_i = (1 - h_i)/(1 - \delta^2(1 - h_i))$, $i = h, c$. The properties of *w* can be derived directly from this expression using the fact that $h_c \leq h_h$. Q.E.D.

Proof of Proposition 3.2

Under a multilateral punishment strategy, the probability that an agent who has ever cheated will be rehired if he cheated or was honest this period and became unemployed is $h_c^c = h_c^h = 0$. The same probabilities for an agent who never cheated are $h_c^h = 0$ and $h_h^h = \tau M/(A - (1 - \tau)M) > 0$. The optimal wage for a cheater is $W^*_c = w(., h_h^c = 0, h_c^c = 0)$, and the optimal wage for an honest agent is $W^*_h = w(., h_h^c > 0, h_c^c = 0)$. Hence because $h_c \leq h_h$ for cheaters and honest agents, proposition 3.1 implies that $W^*_c > W^*_h$. Q.E.D.

Proof of Proposition 3.3

Take the limits of W^* as δ goes to 1 and use the facts that $h_c = h_h = \tau M/(A - (1 - \tau)M)$ under a bilateral punishment strategy and $h_c = 0$ and $h_h = \tau M/(A - (1 - \tau)M)$ under a multilateral

punishment strategy. Use the relations between W^* and the appropriate parameters as specified in proposition 3.1 to take the appropriate limits. Q.E.D.