The Clan and the City:
Sustaining Cooperation in China and Europe*

Avner Greif and Guido Tabellini

Stanford University and Bocconi University

First version: October 2011;

Abstract

Over the last millennium, the clan and the city have been the locus of cooperation in China and Europe respectively. This paper examines – analytically, historically, and empirically – the cultural, social, and institutional co-evolution that led to this bifurcation. We highlight that social organizations, the groups whose members feel moral obligation to, are basic units of cooperation. Social organizations impact institutional development because intra-group moral commitment reduces enforcement cost and social organizations have a comparative advantage in pursuing collective actions. They perpetuate due to positive feedbacks between morality, institutions, and the implied pattern of cooperation.

1 Introduction

Sustaining cooperation is a key challenge for any society. In the course of history, different social arrangements have evolved to cope with this challenge. This paper contrasts the cultural, social, and institutional foundations of cooperation in China and Europe during the last millennium.

In pre-modern Europe, the locus of cooperation were self-governed cities – urban corporations – whose members were drawn from many kinship groups.[1] Cities invested in legal infrastructure, taxed their members and provided them with public goods and social safety nets such as defense, judicial services, education, and poor relief. Despite major economic and social changes, the European city persisted as locus of cooperation to the modern period.

In contrast, clans were the locus of cooperation in pre-modern China since the Sung dynasty (960-1279). A clan is a patrilineal group of related individuals with a common surname traced back to a common ancestor. Clan-based organizations provided public goods and social safety nets - e.g. poor-relief, education, rituals in an ancestors hall, religious services, and protection from non-members, bandits and over-taxation. They also provided daily economic services such as guarding the crops, buying in bulk, maintaining communal burial grounds, and lending money to members in need.[2] The modernization movement in late 19th century China was hostile to the clans that were considered an obstacle to economic development, particularly under the Communist regime. Yet, clans and other clan-like organizations have re-emerged following the reforms in 1978.

How was cooperation sustained in these different organizations? Why did

---

[1]Most European cities west of the Baltic Sea in the north and the Adriatic Sea in the south were self governed. See Pirenne (1969) and Commune (2007).

[2]The discussion of clans in this paper draws on recent surveys and contributions by Haung (1985), Zhenman (2001), Cohn (1990), Szonyi (2002), Faure (2007), Smith (1987); Ebrey and Watson (1986); Freedman (1958), Watson (1982), and Liu (1959). The distinction between the terms 'clan,' 'lineage,' and 'kinship network' notwithstanding (e.g., Watson 1982) we use them inter-changeably.

they become the locus of cooperation in China and Europe respectively? Why did they persist and reemerge despite environmental changes? What imprint did these arrangements leave on current cultural traits in modern Western and Chinese societies? Are they still relevant? These are the questions addressed in this paper.

In any social organization, cooperation is sustained by a combination of intrinsic and extrinsic motivation. Individuals identify with the social group within which they cooperate, and this creates moral obligations. But the decision of joining a group and cooperating is also influenced by material rewards or punishments provided by formal or informal institutions.[3] We view the clan and the city as differing in both respects. In a clan, moral obligations are stronger but are limited in scope, as they apply only toward kin. In a city, moral obligations are generalized towards all citizens irrespective of lineage, but they are weaker, as identification is more difficult in a larger and more heterogeneous group. We refer to this distinction as limited vs generalized morality.[4] Institutional mechanisms also differ between the clan and the city: clan enforcement mainly relies on informal institutions, whereas the city also relies on formal enforcement procedures. In terms of economic efficiency, these two arrangements have clear trade-offs. The clan economizes on enforcement costs, whereas the city exploits economies of scale because it sustains cooperation in a larger and more heterogeneous community.

The first part of the paper investigates the implications of these cultural and institutional differences between clan and city in a theoretical model. We view the evolution of the clan and the city as the result of optimal decisions by individuals with given moral traits. Individuals choose whether to join the clan or the city, based on their moral views and on what others do. A

---

[3]We follow the convention of using the terms 'formal' and 'informal' despite their cultural relativism. By (in)formal institutions we mean institutions that (do not) relay on the coercive power of the state. Most institutions are 'hybrid' of these two institutional types (Greif 2005 and 2006, chapters 1-2).

[4]See Tabellini (2008), Platteau (2000).

first result is that the relative size of the two organizations depends on the distribution of values in society. Large clans are more likely if limited morality prevails, whereas a city is more likely to flourish if moral obligations are not kin based. We then study how morality evolves to reflect prevailing social arrangements. A society in which cooperation occurs within the clan fosters clan loyalty, both in scope and intensity. Conversely, cooperation within a large and heterogeneous population fosters generalized morality and respect for the formal institutions that regulate social interactions in the city. Hence the emergence of one moral system or another is explained by the initial distribution of individuals across organizations, clan vs city.

Combining these two results yields the possibility of cultural and institutional bifurcations. Two otherwise identical societies that differ only in the initial distributions of moral traits evolve along different self-reinforcing trajectories of both cultural traits and organizational forms. Initial diffusion of kin-based morality leads to a steady state where clan loyalty is widespread, the clan is in charge of providing public goods, city size is small and its institutions are weak. This equilibrium captures the arrangements that prevailed in China. Conversely, if generalized morality initially prevails, the organization of society moves to the opposite steady state, where strong and large cities act as the main providers of public goods, as in the evolution of Europe.

Thus, to understand the different paths in China vs Europe, we need to focus on cultural differences in their respective early histories. Even if China and Europe had access to the same technologies, and neglecting the role of geography and other factors, social organizations and cultural traits evolved endogenously and mutually reinforced each other. The second part of the paper compares China and Europe over the last millennium, in light of these theoretical insights.

We focus on two general implications of the theory: first, that clans and cities are alternative ways to organize the provision of public goods; and second, that there is complementarity between these alternative organizational

4

forms and specific cultural traits. Both ideas are supported by the historical analysis. At the turn of the 11th century, China and Europe differed in their cultural traditions and in the strength of kinship based relations, with some of these differences being due to religion. Subsequent legal and institutional developments evolved in different directions in these two parts of the world, strengthening the clan in China and leading to the emergence of strong and self-governed cities in Europe. Moreover, we document the persistence of cultural traits, and how kin based relations remain very strong in China, despite the revolution backlash under the Chinese communist regime and its hostility towards all clan-based organizations.

Cultural traits and organizational forms vary also within Europe and China, and not just across them. In the final part of the paper we exploit this heterogeneity in our empirical analysis. In particular, with regard to Europe, we exploit differences in the early family structures across different parts of Europe, taking family structures as indicators of the scope and strength of kin-based relations. As expected, historical patterns of urbanization reflect these different family traditions, with early urbanization being much more diffused in the European regions that had weak family ties. We also consider differences in clan organization and patterns of urbanization within modern China. As expected, in a sample of Chinese counties, urbanization is negatively correlated with clan strength.

Our explanation of their cultural and institutional bifurcations also contributes to the debate of why Europe and China had distinct development paths. The traditional view, associated with Max Weber, has been that the two societies were structurally different, that Chinese culture was inadequate and led China into a dead end. This view has been challenged as lacking analytical foundations, failing to explain the long periods during which China was economically successful, and wrong in assuming that culture is exogenous and immutable. Weber's opponents have also argued that the rise of Europe was contingent and accidental, and benefited from the flow of ideas,

5

technologies and institutions from the East to West.[5] Our analysis provides analytical foundations to the structural view, highlights that the city and clan have distinct comparative advantages, and recognizes that culture is endogenous and responsive to environmental changes. Yet, it also highlights that complementarity between culture, social organizations, and institutions can lead to persistent cultural and institutional bifurcation.

Finally, the paper also makes a methodological contribution to the analysis of morality, social groups and their role in economic and institutional development. Our approach rests on three central ideas. First, social groups are institutionally important because they constitute basic units of cooperation.[6] Second, groups are also held together by mutual moral obligations and group-specific morality, not just by economic interests or other common attributes. This is what differentiates a local community from a neighborhood, or a nation-state from an empire. To illustrate the last distinction, the Scots and the Irish were subjects of the United Kingdom during World War I but only the former were drafted. The British consider drafting the Irish counter-productive given the lack of moral commitment to the Union (Levi 1997).[7] Third, although very persistent, the forces that glue individuals to a

---

[5]See Deng (2000) for an insightful review. Among the more recent contributions to the debate are Pomeranz (2000), Brenner and Isett (2002), Voigtländer and Voth (2009), Rosenthal and Wong (2011), Goldstone (2008), Ma (2008), Broadberry and Gupta (2006), and Landes (2006).

[6]For a discussion of this duel role of organizations as actors (players) and an institutional element (that change the relevant rules of the game), see Greif (2005, 2006a, 2006 chapters 2 and 5) and Bowles and Gintis (2011).

[7]Groups are seemingly central to the main choice-theoretic approaches concerning morality. In the empirical approach advocated by Guiso et al. (2006), groups are labels assigned to individuals with the same culture. Culture is "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation" (Ibid, p. 23). Yet, 'groups' in this framework are simply labels assigned to individuals with the same culture. Morality itself is sometimes considered a time-invariant characteristic of a society that influences choices over and functioning of formal institutions. For contributions and reviews of the empirical literature, see, for example, Putnam (1993) (on social capital), Guiso, Sapienza, and Zingales (2006); Tabellini (2008); Roland (2004). For contributions and reviews of models of morality, see, for example, Kuran (1987), Rabin (1993), Benabou and Tirole (2006), Bisin and Verdier (2000,

group are not immutable. Individuals tend to vote with their feet and join or abandon a group, and the institutional and cultural foundations of social groups evolve endogenously over time. These central ideas can be used to explore a variety of settings, and not just the institutional development of China and Europe.

The settings we have in mind are those in which extrinsic motivation is too costly to provide without complementary morality. The social organization through which morality perpetuates and to whose members it applies thus matters. Distinct group identities and moralities matter in monitoring workers, violence against women, ethnic violence, terrorism, tax evasion, corruption, and whether the elite invests in promoting general welfare or their control. The role of morality in such phenomena drew scholarly attention but analyses either held morality and/or the moral group as exogenous.[8]

The outline of the paper is as follows. Section 2 develops a model designed to explain the emergence and consolidation of the clan and city. We then apply, in section 3, the resulting theoretical insights to explain the cultural and institutional bifurcation between China and Europe in the distant past. Section 4 presents the empirical analysis exploiting the heterogeneity within China and Europe. Section 5 explains the contemporary revival of the pre-modern cultural and institutional distinctions between China and Europe. Section 6 concludes linking our analysis to these societies' economic accomplishments and challenges.

2   How to Support Cooperation: Clan vs City

This section explores the joint evolution of social organizations and the diffusion of values in society in a choice theoretic framework. The first subsection studies the consolidation of different organizations, the clan vs the city, that

2001), Tabellini (2008a), Bowles and Gintis (2011), and Greif and Tadelis (2010).

[8]For the relevant case of overseas Chinese traders see Landa (1981).

sustain cooperation in the provision of a public good. In a static framework, individuals with given values choose which organization to join (the clan vs the city). The resulting equilibrium configuration depends on the initial distribution of values in society. The second subsection studies how, over time, the diffusion of values is shaped by the prevailing oganizational forms, and explores the joint dynamics of organizations and values.

2.1 Static Analysis

2.1.1 The Model

A population of fixed size M is split in $M > 1$ identical dynasties (or families). Each dynasty contains a continuum of individuals and the size of each dynasty is normalized to unity. Individuals live one period. At the beginning of their life, they choose whether to live in the city or in their clan. There is a single city that can draw inhabitants from all dynasties, while the clan can only draw individuals belonging to a single dynasty; thus, there are several clans and one city. Living in the city (rather than in the clan) gives to each individual an idiosyncratic extra benefit δ, where δ is a random variable distributed within each dynasty according to a uniform distribution over the range $[0, 1/d]$, with $d > 0$.

Each individual has a fixed endowment equal to 1. After having chosen his location (the city or the clan), he decides whether or not to contribute a fixed amount τ to the community where he lives, with $0 < \tau < 1$. Individuals draw utility from two sources (besides the random variable δ defined above). First, they enjoy a material benefit:

$$v = 1 - t + H(g)$$

where $1 - t$ refers to private consumption, with $t = \tau, 0$ depending on the individual choice, and g denotes a public good supplied by the community of residence (the city or the clan). The function $H(.)$ is continuously differ-

entiable, concave and strictly increasing, with $H(0) = 0$ and $H_g(\tau M) > M$ (this last condition implies that the public good is sufficiently productive from a social point of view).

Second, each individual also enjoys a psychological benefit p whenever it contributes a positive amount to the community with which it identifies. Irrespective of where they choose to live, individuals can identify with either their clan or with the city. In each dynasty there is a given fraction of individuals who identify only with the clan, and the remaining fraction identifies with the community where they choose to live, irrespective of whether it is their clan or the city. For shortness, we call them the "clannish types" and the "generalist types" respectively. The psychological benefit of giving to the community with which individuals identify differs by type; specifically, we assume that $p = \lambda$ for a clannish type and $p = \gamma$ for a generalist, with $\lambda > \tau > \gamma > 0$. Thus, identification with the clan is stronger for the clannish type than for the generalist, in the sense that the former draws a more intense psychological reward from participating in the provision of the clan public good. The generalist on the other hand draws the same psychological benefit from contributing to the city or to the clan, but it is a weaker benefit. Note that individuals can only contribute to the community where they live. Moreover, individuals choose where to live, while nature chooses their type. Of course, as described below, identification (or morality) is one determinant of whether individuals choose to live in the clan or in the city. For simplicity, the distribution of the idiosyncratic parameter δ is the same for clannish and generalist types. By assumption, the intrinsic motivation of the clannish types is so strong that they are willing to contribute to their clan (but not to the city) even without any external enforcement ($\lambda > \tau$), while a generalist would choose not to contribute (neither to the clan nor the city) in the absence of enforcement ($\tau > \gamma$).

A simple enforcement technology is available: by spending an amount $e \geq \tau$ of public resources, individuals who do not contribute are detected with

probability q. If caught, their endowment is destroyed, so that their material utility is just equal to $H(g)$ (thus even if caught an individual continues to enjoy the benefits of the public good, or equivalently the public good is non-excludable). The assumption $e \geq \tau$ implies that it would never be optimal to exploit this enforcement technology in the clan, because the clan is too small. If external enforcement is used, it would only be in the city. In this regard, we also assume that:

$$\tau > q \geq \tau - \gamma$$

This implies that, if enforcement is used in the city, it is powerful enough to induce contributions by the generalists, but not by the clannish types. We assume that, whenever it is optimal to do so, the city exploits this enforcement technology.

Table 1 introduces the notation that will be used below, where each cell refers to a combination of preferences and location, where the λ and γ superscripts denote preferences, while the n (for clan) and y (for city) superscripts refer to location.

Table 1

	Location	
Type	*clan*	*city*
clannish	$x^{\lambda n}$	$x^{\lambda y}$
generalist	$x^{\gamma n}$	$x^{\gamma y}$

Thus, $x^{\lambda n}$ is the fraction of individuals who are of type λ (clannish) and live in the clan, $x^{\lambda y}$ the fraction who is of type λ and lives in the city, and so on. We also define $x^{\lambda} = x^{\lambda n} + x^{\lambda y}$, and $x^{\gamma} = x^{\gamma n} + x^{\gamma y} = 1 - x^{\lambda}$; thus, x^{λ} is the given proportion of clannish types in each dynasty, chosen by nature. Moreover, we let $x^{n} = x^{\lambda n} + x^{\gamma n}$ denote the fraction of each dynasty living in the clan, and $x^{y} = x^{\lambda y} + x^{\gamma y} = 1 - x^{n}$ denote the fraction of each dynasty

10

living in the city. The variable x^n is determined in equilibrium, along with the proportions in each cell of Table 1.

With this notation and given the above assumptions, we can pin down public good provision in the two communities. Given that only the clannish types contribute to the clan, the amount of public good provided in the clan is:

$$g^n = \tau x^{\lambda n} \tag{1}$$

Similarly, given that only the generalists contribute to the city, the amount of public good provided in the city is:

$$
\begin{aligned}
g^y &= \tau M x^{\gamma y} - e \ \ \text{if } x^{\gamma y} > e/M\tau \\
g^y &= 0 \ \ \ \text{otherwise}
\end{aligned}
\tag{2}
$$

where we have used the constraint that the enforcement technology can be used only if the city is large enough and it attracts a large enough fraction of generalist types.

Throughout we assume

$$1/2 > e/M\tau \equiv \hat{e} \tag{3}$$

This condition enables the existence of equilibria with small city size (i.e. where less than half the population lives in the city). For reasons that will become clear below, we also assume:

$$\tau M H_g(0) \ < \ 1/d < \lambda + \gamma + q - 2\tau \tag{C1}$$

The timing of events is as follows. Each individual observes his type and then chooses whether to live in his clan or in the city. Individuals then observe the community size (and hence whether the enforcement technology is feasible or not) and choose whether or not to contribute to the community

11

budget. Payoffs are realized. An equilibrium is defined as a distribution of individuals across communities and an allocation of resources such that: i) All individuals have optimally chosen whether to live in their clan or in the city, given their type and correctly anticipating all subsequent events. ii) All individuals have optimally chosen whether or not to contribute to the community budget, given their type, their community size, the distribution of types in the clans and in the city. We only consider symmetric equilibria where all clans have the same size and the same distribution of types.

2.1.2 Equilibrium city size

We now discuss how different types are distributed amongst the city or the clans. In general, many equilibria are possible, depending on parameter values, and multiple equilibria can also occur. In particular, we cannot rule out equilibria in which the whole population lives in the same community (city or clans). The reason is that there are economies of scale in public good provision. Thus, if I expect everyone else to live in the city (or in the clan), then it is also optimal for me to do so.

In this subsection we characterize the possible equilibria, paying particular attention to equilibria where city and clans coexist. We need some additional notation. Let $z = n, y$ denote the community of choice, namely the clan (n) or the city (y). As before, let $p = \lambda, \gamma$ denote the preference type, namely clannish (λ) or generalist (γ). Expected equilibrium welfare of type p in community z is a known function of the composition of the population living in community z, namely: $W^{pz}(x^{\lambda z}, x^{\gamma z})$, where $x^{\lambda z}$ denotes the fraction of λ types in each dynasty who have chosen community z, and $x^{\gamma z}$ denotes the fraction of γ types in each dynasty who have chosen that same community - cf. Table 1.

Let δ^p denote the value of δ that leaves type p indifferent between the clan and the city, for $p = \lambda, \gamma$. That is, δ^p is obtained from (9) setting $W^{pn} = W^{py}$ and hence it is a known function of x^{pn}, x^{py}, $p = \lambda, \gamma$. the appendix proves:

Lemma 1 $\delta^\lambda = \delta^\gamma + a, \quad where \ a \equiv \lambda + \gamma + q - 2\tau > 0$

The last inequality follows from previous assumptions about parameter values. It implies that the attractiveness of the clan compared to the city is always higher for the clannish types than for the generalists, given that they have the same realization of δ, for all possible distribution of players across communities. This result is important, because it has implications about the relative distribution of types across communities in equilibrium. The condition $\lambda > \tau$ that gives rise to $a > 0$ means that the amount to be contributed is small relative to the intrinsic reward from not cheating in the clan. This implies that the temptation to free ride is not so strong that the clannish types are easily attracted to the city where they can free ride on the generalists.

This notation enables us to express the fraction of individuals of each type that prefer to be in the city or clan respectively as:

$$x^{pn} = \Pr(\delta \leq \delta^p)x^p \tag{4}$$

$$x^{py} = \Pr(\delta > \delta^p)x^p = 1 - x^{pn} \tag{5}$$

Given the assumed distribution of δ, we thus have:

$$
\begin{aligned}
x^{pn} &= d\delta^p x^p \ \text{if} \ \delta^p \in [0, 1/d], \\
x^{pn} &= 0 \qquad \text{if} \ \delta^p \leq 0, \\
x^{pn} &= x^p \qquad \text{if} \ \delta^p \geq 1/d
\end{aligned}
\tag{6}
$$

and correspondingly for x^{py}.

There are five possible kinds of equilibria, and some of them can exist under the same configuration of parameter values (i.e. we cannot rule out multiple equilibria). In each equilibrium, city size (and hence clan size) is given by a different expression.

(i) First, as already stated, we could have an equilibrium in which every-

one joins their clan. By Lemma 1 this requires $\delta^\gamma \geq 1/d$. Here city size is 0.

(ii) Second, we could have the opposite situation, where everyone is in the city. By Lemma 1, this requires $\delta^\lambda \leq 0$. Here city size is the full population.

(iii) Third, we could have an equilibrium with full sorting, where all the clannish types are in the clan, while all the generalists are in the city. This requires $\delta^\lambda \geq 1/d$ and $\delta^\gamma \leq 0$. This too could be an equilibrium, for instance if λ is large enough and $1/d$ is small enough. Here city size is Mx^γ, namely it coincides with the generalists' population.

(iv) Fourth, we could have an equilibrium with segregation, where all the clannish types are segregated in the clan while the generalists join both communities. This requires $\delta^\lambda \geq 1/d$ and $1/d > \delta^\gamma \geq 0$. Here, city size is $Mx^{\gamma y}$ (the fraction of generalists living in the city), which in turn is pinned down by the generalist types' indifference condition.

(v) Fifth, we could have the opposite equilibrium with segregation, namely the generalists are segregated in the city, while the clannish types join both locations. This requires $0 < \delta^\lambda < 1/d$ and $\delta^\gamma \leq 0$. Here, city size is proportional to $x^\gamma + x^{\lambda y}$, and $x^{\lambda y}$ is pinned down by the clannish types' indifference condition.

Note that Lemma 1 rules out an equilibrium with full sorting where all the clannish types prefer to be in the city, while the generalists all prefer to be in the clan (i.e. the opposite of equilibrium (iii)). Moreover, Lemma 1 plus the second inequality in (C1) rule out the equilibrium in which both types are present in both the clan and the city.[9]

It is tempting to interpret these admissible equilibria as reflecting a historical transition that accompanies the growth in the size of cities. On can think of an early situation where everyone lived in a clan-based society (equilibrium (i)). Then, for a variety of exogenous reasons (trade opportunities,

---

[9]In an equilibrium in which both types are present in both locations we should have $1/d > \delta^p > 0$, for $p = \lambda, \gamma$. But by Lemma 1 this is impossible if the second inequality in condition (C1) holds.

14

or defense) the city becomes more appealing (the distribution of the random variable δ shifts to the right), and some - but not all - generalist types are attracted to the city, so that we move to equilibrium (iv). As the city becomes even more attractive (as the distribution of δ keeps moving right), all the generalist types move to the city, and we shift to the equilibrium with full sorting (equilibrium (iii)). If the attractiveness of the city rises further, eventually the city becomes appealing even for the clannish types, and we shift to equilibrium (v), until the last step where all the clannish types have moved to the city and we are in equilibrium (ii). This interpretation has many loose ends, however. First, because everything occurs in reaction to a change in an exogenous parameter (the distribution of δ) on which the theory has nothing to say. Second, because the theory is silent about equilibrium selection (i.e., we cannot explain how we move from one equilibrium to another).

Rather than pursuing this interpretation, therefore, in the next section we introduce an explicit source of equilibrium dynamics through changing preferences. We then study how changing preferences interact with the distribution of individuals across communities, within the same equilibrium.

In preparation for that, here we ask how existence of the equilibrium and equilibrium city size depend on the preference composition of the population (the parameter x^λ in the model). In the first two kinds of equilibria, the answer is immediate. In these two equilibria the whole population is concentrated in a single location (the clan or the city), and thus city size cannot depend on the composition of the population. Moreover, as discussed above, these equilibria always exist for large configurations of parameter values and for any preference composition of the population, as captured by x^λ.

The remaining three equilibria (with full sorting and segregation) only exist for some values of the fraction of the clannish types within each dynasty, x^λ. Each equilibrium exists if and only if x^λ belongs to a specific sub-interval of $[0, 1]$. Under conditions stated in the appendix, there are three contiguous and non-overlapping sub-intervals of $[0, 1]$ such that, as x^λ moves from the

15

lowest to the highest sub-interval, the equilibrium shifts from segregation in the city, to full sorting, to segregation in the clan. More precisely, the appendix proves:

Proposition 2 *Under conditions (A1-A4) in the appendix:*

(i) The equilibrium with segregation in the city exists if and only if $x^\lambda \in (x^{\min}, \underline{x})$, where $1/2 > \underline{x} > x^{\min} \geq 0$

(ii) The equilibrium with full sorting exists if and only if $x^\lambda \in [\underline{x}, \bar{x}]$, where $1 > \bar{x} > 1/2 > \underline{x} > 0$

(iii) The equilibrium with segregation in the clan exists if and only if $x^\lambda \in (\bar{x}, x^{\max})$, where $1 - \hat{e} > x^{\max} > \bar{x}$.

This Proposition is depicted in Figure 1, and the thresholds of each interval are implicitly defined in the appendix.

<<insert figure 1 around here>>

Finally, the appendix also proves that, in all three equilibria and if the first inequality in condition (C1) above holds, equilibrium city size decreases with x^λ, the fraction of clannish types within each dynasty. The intuition is simple. Consider first the equilibrium with full sorting. Here the result is immediate: since all the clannish types live in the clan while all the generalists live in the city, an increase in x^λ shrinks the fraction of generalists and hence city size shrinks too. Next, consider the equilibria with segregation. As the fraction of clannish types increases, the clan becomes more attractive compared to the city, because the amount of public goods increases in the clan or it shrinks in the city (recall that in equilibrium clannish types contribute to the public good in the clan but not in the city, while the generalists do the opposite). Hence, as x^λ increases, some individuals who were close to indifference move from the city to the clan - which makes the clan even more attractive because it can better exploit the economies of scale in public goods provision compared to the city. The first inequality in condition (C1)

16

is needed to make sure that, as this happens, equilibrium is restored because the distribution of the idiosyncratic preference parameter δ is sufficiently spread out.

Note that, as we shift from equilibrium (i) in Proposition 1, to equilibrium (ii), to equilibrium (iii), city size increases. We thus have that, as the fraction of clannish types in the population increases, equilibrium city size shrinks (and clan size correspondingly increases), either as we shift from one equilibrium to the other, or within the same equilibrium.

2.2 Dynamics

The previous subsection pointed out how the distribution of values in the population affects the relative size of the clan vs the city. In this section we discuss the opposite link: how the distribution of individuals between the clan and the city has implications for the endogenous evolution of values in society. We then discuss the two-way interaction between endogenous values and the relative size of the city vs clan.

2.2.1 The Dynamic Model

The parameter x^λ, measuring the fraction of clannish types within in each dynasty, captures the distribution of values in society. Here we assume that this parameter is not fixed, but varies period after period reflecting the initial distribution of individuals between the clan and the city.

Suppose that the same environment described in the static model reproduces itself in each period with a constant population of one-period lived individuals. We can interpret individuals like households, and assume that each individual (or household) gives birth to a new individual (or household). The preferences (λ vs γ) of the newborn individual reflect both the preferences his parent, as well as the community (clan vs city) where his parent lived. Specifically, we assume that a clannish parent living in the clan gives

17

birth to a clannish son with probability $\bar{p} \geq 1/2$, and to a generalist son with probability $1 - \bar{p}$. But a clannish parent living in the city gives rise to a clannish son with probability $\underline{p} < \bar{p}$, and to a generalist son with probability 1-$\underline{p}$. By symmetry, a generalist parent gives birth to a generalist son with probability $\bar{p}$ if he lives in the city, and with probability $\underline{p}$ if he lives in the clan, giving birth to a clannish son with complementary probabilities. These assumptions are meant to capture the idea that values are transmitted both within the family, as well as by the environment where one grows up. Parents are likely to transmit their values to their children, either deliberately or by setting an example - hence the assumption that $\bar{p} \geq 1/2$. But at the same time, growing up in the clan is likely to foster loyalty to and identification with the clan, while growing up in the city is likely to foster a more generalized respect for the rule of law, which in our setting is captured by the formulation of generalist preferences - hence the assumption that $\bar{p} > \underline{p}$. We also assume that $(\bar{p}+\underline{p})/2 > 1/2$, namely on average a clannish parent is still more likely to give rise to a clannish son than to a generalist, irrespective of where he lives. This assumption can be interpreted as saying that the influence of the family is stronger than that of the environment where one grew up. Finally, we assume that the distribution of the idiosyncratic δ parameter remains the same within each group of clannish and generalist types, irrespective of the relative size of the two groups.

Under these assumptions, the law of motion of the fraction of clannish types within each dynasty is (period $t + 1$ refers to the newborn generation, while period t refers to the parents' generation):

$$x_{t+1}^{\lambda} = \bar{p}\, x_t^{\lambda n} + \underline{p}\, x_t^{\lambda y} + (1 - \bar{p})\, x^{\gamma y} + (1 - \underline{p})\, x_t^{\gamma n} \tag{7}$$

The terms on the right hand side of (7) refer to the fraction of clannish sons born respectively from: clannish parents living in the clan; clannish parents living in the city; generalist parents living in the city; generalist parents living in the clan. Recalling that $x_t^{\lambda} = x_t^{\lambda n} + x_t^{\lambda y}$ and that $x_t^n = x_t^{\lambda n} + x_t^{\gamma n}$, we can

rewrite (7) as:

$$x_{t+1}^\lambda = (1 - \bar{p}) + (\bar{p} + \underline{p} - 1)\, x_t^\lambda + (\bar{p} - \underline{p})\, x_t^n \qquad (8)$$

We can interpret equation (8) as follows: the first term on the right hand side, $(1 - \bar{p})$, is the birth rate of clannish types that would occur if the previous generation only consisted of generalists living in the city. The second term is the differential birth rate between clannish and generalist parents, $(\bar{p} + \underline{p} - 1)$, times the fraction of clannish parents. And the third term, $(\bar{p} - \underline{p})$, is the differential birth rate between parents living in the clan and the city, times the fraction of parents living in the clan.

Equation (8) can be combined with the results in the previous subsection to obtain a full dynamic analysis. In equilibrium, the composition of types within each dynasty is jointly determined with the allocation of individuals between the clan and the city. Since different equilibria are possible in the static part of the model, we have to consider each of them in turn. We neglect the trivial equilibria in which all the population is in a single location (city or clan), focusing instead on three static equilibria: the one with full sorting of types across communities, and the two equilibria with segregation of one type in one community.

2.2.2 Dynamic Equilibrium with Full Sorting

Consider first the equilibrium with full sorting of types across communities. This is simple, because the fraction of each dynasty living in the city is just $x_t^y = 1 - x_t^\lambda$ - cf. (14) in the appendix. Hence, (8) reduces to:

$$x_{t+1}^\lambda = (1 - \bar{p}) + (2\bar{p} - 1)\, x_t^\lambda$$

Hence, this dynamic equation has a single steady state (denoted with an s subscript), $x_s^\lambda = 1/2$. Not surprisingly, given the symmetry of the model, in the steady state the population is split in half, with all the generalists living

in the city and all the clannish types living in the clan. Moreover, since $1 > \bar{p} \geq 1/2$, the steady state is stable and the adjustment to the steady state is monotonic. While we remain in this equilibrium, any permanent change in any of the parameters of the model has no effect (either temporary nor permanent) on the preference composition of the population, nor on the distribution of types across localities. Intuitively, with full sorting, the distribution of individuals across communities is entirely driven by their preferences, and cannot be affected by other parameters of the model. Hence the evolution of preferences in society is also entirely determined by its own past history and cannot reflect the influence of any other economic or social force.

By the results of the previous subsection, we have an equilibrium with full sorting if and only if $x_t^\lambda \in [\underline{x}, \bar{x}]$. Combining these dynamic results with those of the previous subsection we thus have:

Proposition 3 *Suppose that the initial fraction of the clannish population, x_0^λ, is such that $x_0^\lambda \in [\underline{x}, \bar{x}]$. Then ,overtime the fraction of population with clannish values converges to $x_s^\lambda = 1/2$, and both in the steady state and throughout the adjustment process all the clannish types live in the clan while all the generalists live in the city.*

2.2.3   Dynamic Equilibrium with Segregation in the Clan

Next, consider the equilibrium where the clannish types are segregated in the clan, while the generalists are present in both the clan and the city. Here the derivation is more cumbersome, so we relegate it to the appendix. But the logic is simple. In this equilibrium, some generalists are attracted to the clan. This in turn influences the preferences of their offspring, which are more likely to become clannish types. Hence in the steady state the clannish population exceeds 1/2. If the differential birth rate of clannish types between parents living in the clan and in the city (the term $(\bar{p} - \underline{p})$) is small, however, the steady state is stable. Hence, given that the initial condition of preferences

is in the region corresponding to this equilibrium, society remains forever in this region and it converges to a steady state where the clannish types are a majority and city size is correspondingly small.

More precisely, let $x^{\lambda c}$ denote the steady state fraction of clannish types corresponding to this equilibrium. The appendix proves:

Proposition 4 *Suppose that the initial fraction of the clannish population, x_0^{λ}, is such that $x_0^{\lambda} \in (\bar{x}, x^{\max})$. Suppose further that $x^{\lambda c} \in (\bar{x}, x^{\max})$ and that $(\bar{p} - \underline{p})$ is sufficiently small. Then, overtime the fraction of population with clannish values converges monotonically to $x^{\lambda c} > 1/2$, and both in the steady state and throughout the adjustment process all the clannish types live in the clan while the generalists mix between the clan and the city.*

As further discussed below, this steady state, with most of the population living in the clan and only a minority in the city, captures the social organizations that prevailed in China - hence the c superscript to denote this equilibrium.

Note that in this equilibrium, changes in the deep parameters of the model have permanent effects on the distribution of values in society - they impact on the steady state $x^{\lambda c}$. For instance, if the clan becomes more efficient in providing public goods to its members, more generalists are attracted to the clan, and overtime a larger fraction of the population acquires clannish values.

2.2.4 Dynamic Equilibrium with Segregation in the City

Finally, consider the other equilibrium with segregation, where the generalist types are all in the city, while the clannish types are present in both communities. Here the logic is the reverse of that in the previous equilibrium. Since some clannish types are attracted to the city, their offspring are more likely to become generalists. Hence in the steady state the fraction of clannish types in the population ends up being a minority, and if $(\bar{p} - \underline{p})$ is small the

steady state is stable. More precisely, let $x^{\lambda e}$ denote the steady state fraction of clannish types corresponding to this equilibrium. The appendix proves:

Proposition 5 *Suppose that the initial fraction of the clannish population, x_0^{λ}, is such that $x_0^{\lambda} \in (x^{\min},\underline{x})$. Suppose further that $x^{\lambda e} \in (x^{\min},\underline{x})$ and that $(\bar{p} - \underline{p})$ is sufficiently small. Then, overtime the fraction of population with clannish values converges monotonically to $x^{\lambda e} < 1/2$, and both in the steady state and throughout the adjustment process all the generalist types live in the city while the clannish mix between the clan and the city.*

This steady state, with most of the population living in the city and only a minority in the clan, captures the social organizations that prevailed in Europe - hence the e superscript to denote this equilibrium.

Here too, as in the previous equilibrium, the steady state distribution of values is affected by changes in the model's parameter: whatever makes the city more or less attractive to the clannish types has permanent effects on the proportion of individuals with clannish values.

2.3 Discussion

Contrasting the last three propositions, we see that even small differences in the initial social and moral conditions in an otherwise identical economic and social environment lead to lasting and marked distinctions in both the distribution of values and the organization of society. Moreover, changes in the institutional underpinnings of the clan vs the city, by making these two organizations more or less attractive, have permanent effects on citizens values. These permanent effects reflect the complementarity between morality and the arrangements to provide public goods.

A society that starts out with a diffuse sense of loyalty to the clan will find it optimal to mainly rely on the clan to provide public goods, and only a small fraction of the population will be drawn towards the city. This situation corresponds to the equilibrium with segregation in the clan, where

all the clannish types remain in the clan, and the generalists are distributed both in the clan and in the city. In this equilibrium city size is small, and clan size is large, for two reasons: first, because the majority of the population has clannish values, and all these individuals find it optimal to remain in the clan; second, because only some of the individuals with generalized morality (the generalists) are induced to stay in the city (these are the individuals with a high realization of the δ parameter). Being large, the clan is more attractive than the city, since even a clan can exploit the economies of scale associated with public good provision. Moreover, the clan can also rely on many loyal individuals whose values make the clan work smoothly with no free riding. This situation preserves itself overtime, because the generalists who are attracted to the clan are more likely to give rise to clannish off-spring, compared to the generalists who live in the city, further reinforcing the comparative advantage of the clan over the city. This equilibrium is thus meant to capture the arrangements that prevailed in China, where loyalty to the clan was widespread and the clan was the main organization in charge of providing public goods.

Conversely, a society that starts out with a widespread diffusion of values consistent with generalized morality finds itself in the opposite situation. As described in the equilibrium with segregation in the city, here all the generalists remain in the city, while the clannish types are found in both locations. Here city size is large (and clan size is small), for two reasons. First, since there are many generalists, the city operates smoothly and with little free riding; second, a large city provides a large quantity of the public good compared to the clan. Both reasons make the city attractive compared to the clan. Again, this situation preserves or strengthens itself over time, as the clannish types who are attracted to the city are more likely to give birth to generalist types. This equilibrium is meant to capture the evolution of Europe, where clan loyalty was replaced by generalized morality and where the city became the main arrangement to provide public goods.

These theoretical results draw attention to the key complementarity between culture and the organization of society in the provision of public goods. To understand why social institutions evolved along different paths in China vs Europe, we need to focus on cultural differences in their respective early histories. Even if China and Europe had access to the same technologies, and even neglecting the role of geography and other features of the environment, endogenous social institutions and individual cultural traits mutually reinforced each other. Different initial conditions in the diffusion of specific cultural traits can account why social institutions and morality evolved in different directions in these two parts of the world.

3   History

Our formal analysis reveals how two otherwise identical societies that differ only in the initial distribution of values can evolve along different self-reinforcing trajectories of cultural traits, organizational forms, and institutions.

This section evaluates this result based on the historical experience of China and Europe. We first discuss why around the year 1000 AD generalized morality was more prevalent and kinship groups were less prevalent in Europe and why, at that time, people in China and Europe could self-organize themselves. We then establish that both systems evolved over time through positive feedbacks among social organization, morality, and institutions. We particularly focus on evaluating the importance of unobserved moral distinctions. Finally, we transcend the boundaries of our model by incorporating the state in the analysis. The evidence concerning the integration of clans and city in the administrations of the Chinese empire and European states confirms their importance and highlights the positive feedbacks between social organizations and the organization of the state.

3.1 Initial Conditions

Circa the beginning of the second millennium individuals in China and Europe had much discretion in forming new social organizations. In each region, however, they made this choice facing different initial conditions. In China, large kinship organizations were common while in Europe they were rare. The corresponding moral traditions, Confucianism in China and Christianity in Europe respectively emphasized moral obligations among kin and general moral obligations. These different initial conditions reflect political and religious processes exogenous to the dynamics we examine and thus can be taken as predetermined in our analysis.

3.1.1 China

In China, the Han dynasty (206 BCE – 220 CE) came to power advocating Confucianism as an alternative to the Legalism of the previous Qin dynasty. While Legalism emphasizes legal obligations, Confucianism considers moral obligations among kin as the basis for social order. Confucius placed obligations to kin above obligations to the law. His Analects contains the following conversation. "The Governor ... said to Confucius, 'In our village there is a man... when his father stole a sheep, he gave evidence against him.' Confucius answered 'In our village those who are straight are quite different. Fathers cover up for their sons, and sons cover up for their father. Straightness is to be found in such behavior" (XIII, 18). By the time the Han dynasty collapsed, however, Buddhism spread in China perhaps because its spirituality offered comfort in these difficult times. In any case, Buddhism had challenged Confucianism and, in particular, had undermined large kinship organizations and limited morality by emphasizing the individual, monasticism, and the religious community.

Buddhism was particularly promoted by the many non-ethnic Chinese rulers of the various states that emerged in China at the time because it transcends ethnic identities. The ethnically-Chinese Tang emperors (618-

25

907) that reunified China initially also supported Buddhism. Eventually, however, they turned against it to gain political legitimacy, expropriate the wealth of the Buddhist organizations and subject their inhabitants to taxation.[10] With the support of the state, Confucian scholars had responded to the popularity of Buddhism by formulating the so called Neo-Confucianism. It synthesizes Taoist cosmology and Buddhist spirituality around the core of Confucian concern with kinship. Relations among kin thus remained ideologically salient. Buddhist scholars responded in kind by emphasizing filial obligations.

3.1.2 Europe

In Europe, the Germanic invasion of the Roman Empire initially reinforced tribalism but by the 9th century the nuclear family predominated. Large kinship groups remained only on Europe's social and geographical margins (e.g., high nobility and Scotland).[11] Clearly, the European family structures did not evolve monotonically toward the nuclear family, nor was their evolution geographically and socially uniform. Yet, the evidence reflects a general process of decline in the importance of large-kinship groups. For example, the (Germanic) Salic law of the sixth century, denied legal rights to anyone not affiliated with a large kinship group. Yet, by the eighth century

---

[10]To illustrate the penetration of Buddhism note that in 842 alone, the property of more than 4,600 Buddhist places of worship (with 415,000 inhabitants) was confiscated and their inhabitants were put on the tax-role. By the 960, only 2,694 monasteries out of 30,336 had remained (Kuhn 2009, pp. 14-5).

[11]E.g., Mitterauer et. al. (1982); Goody (1983); Ekelund et. al. (1996); Herlihy (1985); Greif (2006), chapter 8. Quantitative evidence from later centuries also reveals a decline of large kinship groups. English court rolls from the thirteenth century reflect that cousins were not more likely than non-kin to be in each other's presence (Razi 1993). The English poll-tax records of 1377 indicate that there were approximately 2.3 individuals over the age of thirteen per-household (Schofield 2003, p. 83). The mean household size in five English parishes in the 16th century ranged from 4.05 to 6.05 and that of 100 parishes from the 16th to the 19th century was 4.788. Only about 10 percent of the households had a resident kin (Laslett 1969, pp. 204, 207, 218). The mean household size of those receiving poor relief in Strasbourg in 1523 was similar (Jutte 1996, p. 382).

the term family among the Germanic tribes denoted one's immediate family. By the 10th century, the English law mandated every male to join a group that would guarantee his appearance in court, suggesting that kinship groups could no longer be held accountable as was the case when the Salic law was specified. Tribes and lineages, by and large, were no longer institutionally relevant (Guichard and Cuvillier 1996).

Christianity fostered the dissolution of kinship groups in Europe as it emphasized the salvation of the individual, the community of believers, and generalized morality. By the 4th century, the Church began advancing a marriage dogma that prohibited practices that sustain kinship groups, such as polygamy, concubinage, marriages among distant kin; moreover, marriages without the woman's consent were discouraged.[12]Goody (1983) has conjectured that this was a self-serving dogma and Ekelund et al. (1996) articulated on the idea that the "Church hoped that by further restricting marital options, the incidence of heirs would decrease, thereby increasing the likelihood of inheritance for itself" (pp. 93-4).

Summarizing: the political and religious processes that shaped social organization and morality in Europe and China had caused them to differ circa 1000. In China, kinship relations beyond the immediate family were salient and obligations toward kin were stressed, while in Europe such relations were relatively muted and generalized morality was stressed.

3.2 The Outcomes

Consistent with our analysis, bifurcation is the hallmark of the subsequent evolution of social organizations, morality, institutions and cooperation in China and Europe.

---

[12]See Goody (1983), Ekelund et al (1996) and Greif 2006a.

3.2.1 Social Organizations: Clans vs Cities

For about three of for centuries, beginning around 1000 CE, both the Europeans and Chinese expanded into new areas gained through migration, reclamation and conquest.[13] On both regions individuals and groups, rather than the state, create most of the new settlements. Our analysis implies that when individuals choose their social affiliation, the impact of morality on their choices should be particularly pronounced. In fact, the new communities differed as predicted by our analysis if initial distinct moralities prevailed. In particular, Chinese settlements were clan-based while the European ones were not.

To present the case of China consider figure 2 that presents temporal and spatial distributions of lineage genealogy roles. The data is far from perfect and the selection bias is unknown to us. It is nevertheless revealing. The figure presents the data for each of China's main regions and to the end of each dynasty (republic): 907 (Tang), 1297 (Sung), 1368 (Yuan), 1648 (Ming), Qing (1911), and the republic (1949).

<< Insert Figure 2 around here >>

During the Tang dynasty, the North was the core of the Empire and note that there were no clan there (according to our measure) to 1648. Clans were not the initial condition. In the 10th century, however, a period of cold climate (the Third Little Ice Age) had begun and negatively impacted agricultural production in north China. Moreover, the region was conquered (in 1127) by the cavalry-based army of the Tungusic Jurchen. The Chinese began to migrate to the rice growing area of the Lower Yangtze. Rice cultivation was less effected by the climatic change and the rice paddies and the waterways was advantageous to the Chinese infantry. his migration and

---

[13]By the late 14th century, as the south was settled, the area under cultivation increased from 255 to 522 million mu (Chao 1986, p. 89). Cultivated land almost double from 1400 to 1650 (Maddison 2007; Meyer and Wang 2002, p. 571).

settlement process was generally uncoordinated by the authorities and took place in areas with weak or non-existing state administration. The political conditions were conducive for people to self-organize themselves in establishing new settlements (Kuhn 2009, chapter 4).

Unlike the villages in the North, the new communities had stronger clan structure as evident from leaving lineage genealogy behind. Settlement was "based on kinship ties" and migrants "constructed a new kin-group on the frontier for the purpose of land clearance and developing an irrigation infrastructure" (Rowe 2002, p. 534). The territorial expansion move further south particularly after the Mongol (Yuan) invasion in the 13th century. Clan genealogies mark the trail. In north China, large-scale landlords emerged following the large decline in population following the Manchurian (Qing) invasion of the 17th century. The need for better protection fostered clan-based rural organizations while townships became the residence of large landlords and thus the center of the locally powerful clans (Skinner 2001). The west (Upper Yangtze) was rapidly populated, and clan genealogies survived following the military expansion under the early Qing.

Clans remained important economic actors to the modern period. While data is incomplete, as late as the turn of the twentieth century between 40 to 87 percent of the villages in the south were composed of members of one clan while in the north multi-clans villages with clan-based governance predominated (Haung 1985). Most of the about 50 Chinese firms that registered in Hong Kong from 1920 to 1937 were probably "closely related to Chinese tongs [clan trusts] organized along familial and regional lines" (Chung 2010, p. 1426). In each of at least 21 companies, the directors and shareholders had the same surnames and the same places of origin and some even shared the same registered addresses. Similarly, in the various regions in the Fujian province about 50 percent of the 'feudal land' (defined as land not personally cultivated by its owner) was owned in common by lineages (figure 3).

<< figure 3 around here >

In Europe, following the mid 9th century, the political situation was also conducive to the creation of new self-governed organizations. Attacks by Muslims, Vikings, and Hungarians weakened the secular authorities while the Papacy was captured by the Roman nobility and was in no position to lead. Moreover, advances in metallurgy and the decline in superstition fostered land clearing. These factors provided incentives and opportunity for Europeans to self-organize themselves in creating new settlements. The area under cultivation increased by a factor of six between 950 and 1350 (William 1999, p. 39).

The patterns of new settlements in Europe substantially differed. Since the late medieval period, communities were composed of non-kin and the resulting urban network still prevails (Clark 2009). Subsequent immigration was generally encouraged, newcomers had legal rights and citizenship could be acquired based on such factors as wealth, tax payments, length of residency, and religion.[14] To illustrate, more than 1500 towns were created east of the Elbe (Cantoni and Yuchtman 2009) and the number of chartered cities rose in England by the factor of 12 from 1066 to 1600 (figure 4a).[15] Similarly, there was a large increase in the number of European cities with population of at least 10,000 people (figure 4b).

<< Insert figures 4a, 4b here >>

There were, obviously, cities in China. Yet, our analysis predicts that intra-clan loyalty would limit urbanization. We further evaluate this prediction in section 6. Here we only note (table 2) that China's urbanization rate remained between three to four percent from the 11th to the 19th century while the initially lower urbanization rate in Europe rose to about ten percent. Including small cities urbanization rates were comparable, but these

---

[14]E.g., By 1245, foreign university scholars in Bologna gained the same civil rights as Bolognese citizens although without the right to vote or to hold office (Clark 1987).

[15]E.g., Clark (2009) and Beresford and Finberg (1973).

small cities (townships) were venues for cooperation among members of local clans rather then their melting pot (Skinner 2001).

Table 2: Percentage of Population in Cities of 10,000 or More.

	Year			
	1000	*1500*	*1650*	*1800*
China	3.1	3.8	4	3.8
Europe	0	5.6	8.3	10

Source: Table 1.8 on page 40 in Maddison. 2007.

China and Europe characteristic social organizations – the Clan and the City – emerged through a process of self-selection influenced by initial cultural and social conditions. They perpetuated and proliferated since. On the one hand, "the clan as a Chinese institution is the pre-modern period is generally believed to have prevailed some 800 year, beginning with the Sung dynasty" (Fei and Liu 1982, p. 393).[16] On the other hand the period from 950 to 1350 witnessed "the birth of cities" in Europe (Pirenne 1969, p. 213).

3.2.2 Enforcement and governance

Our analysis reveals why formal enforcement institutions are likely to be necessary in the city but not in the clan. Cooperation in the European cities and the Chinese clans was motivated differently.

European cities increasingly developed formal, legal institutions to support cooperation among the individuals they attracted from the countryside. There was a transitions from 'hand-shakes' to contracts and from electing voluntary judges who used customary law to professional judges relying on a formal legal codes.[17] The first, post-Roman law school in Europe was established in the late 11th century in the University of Bologna (Italy) and

---

[16]Ebrey and Watson, (1986), Watson (1982), Fei and Liu (1982) p. 393.

[17]Cantoni and Yuchtman (2009) provide an illuminating quantitative analysis of the impact of legal professionals on late medieval trade.

by the 13th century it had about 1,000 students. Many more students were attending the more than 46 additional universities that were established by then (Clark 1987).

Clearly, cooperation in Europe was also based on generalized morality (and private-order institutions). The role of moral commitment in fulfilling contractual obligations is suggested, for example, by contracts that could not be legally enforced such as creating self-governed cities and limiting how much an oversea agent would defraud a faraway merchant. Similarly, organizations (such as guilds) provided club-goods and fostered cooperation among non-kin by the threat of exclusion. Intra-city enforcement institutions and generalized morality facilitated intra-city, inter-lineage cooperation in confronting feudal lords and mitigating internal conflicts.

In China, intra-clan dispute resolution mechanisms were less formal then intra-city ones although clans had rule-books and adjudication procedures. Inter-clan legal disputes were arbitrated by the clans' elders and compromise was the main principle of adjudication. More generally, compared to Europe the Chinese legal system was based more on custom and pragmatism than on the law and judicial principles. In terms of our analysis, limited morality reduced intra-clan enforcement cost, thus reducing the potential gains from creating a more formal legal institutions.[18].

While the European cities relied on taxes and monopolies for revenue, Chinese clans relied on voluntary contributions to finance their activities. The Chinese 'clan trust' was first introduced during the Sung Dynasty and it enabled clan members to jointly hold property. Trusts were endowed by wealthy clan members and some clans, particularly in the south were very wealthy. In the north, lineage organizations had little, if any, property and their operation was financed by on-going contributions. In multi-clan villages the local temple collected contributions and assisted members of the local clans (Huang 1985).

---

[18]See Ma (2007) for a review. Cf. Zelin et al (2004).

The internal organization of cities in China and Europe is consistent with the prevalence of limited and generalized morality respectively. Limited morality increases the cost of cooperation across clan boundaries implying that if it prevailed in China, intra-city public good provision should have been provided by either the state or by each clan to its members. Indeed, Chinese cities were governed by government-appointed officials and residential areas were divided into smaller wards, called fang, each of which was surrounded with walls and guarded gates (e.g., Huang 2006).[19] This urban structure reinforced limited morality because each fang was usually dominated by people from the same clan or several clans from the same region (Rowe 1984, Huang 2006).

Most European cities in the West, but no Chinese city, had gained self governance by 1500.[20] The lack of self-governed cities in China was not simply due to a more powerful state, but also due to pervasive kinship structure that facilitated state control over cities. Indeed, immigrants to cities remained affiliated with their rural kinship groups. As late as the 17th century, in a relatively new city "the majority of a city's population consisted of so-called sojourners, people who had come from elsewhere and were considered (and thought of themselves as) only temporary residents .... suspicions were always rife that sojourners could not be trusted" (John Friedmann 2007, p. 274).

General morality fosters interest-based associations among non-kin while limited morality fosters kin-based associations. The comparison between Europe and China is striking. In Italy alone 1385 guilds were created between

---

[19]Capital cities, in addition, were divided into five wards whose officials were responsible for security and fire prevention

[20]Late medieval Germany (1939 borders) had 2,256 incorporated cities (Cantoni and Yuchtman 2009). The cities of Flanders were self-governed since the twelfth century (Nicholas 1992, pp. 119-23). Twenty percent of the population of Castile (in 1500) lived in its 30 self-governed cities (Nader 1990, p. 3). Most French cities had at least limited feudal obligations (Medieval Commune 1911) and there were 500 charted English cities by 1500 (Ballard 1913; Ballard and Tait 1923; Beresford and Finberg 1973).

1220 and 1800 (Mocarelli 2008, table 1). In China, from the Song dynasty (960-1279) and well into the Ming dynasty (1368-1644) occupation groups in a city had "headmen" (*hangtou* or *hanglao*). These, however, were subservient to the authorities and functioned mainly as price setters and brokers.

By the late Ming, self-governed guilds (*huiguan*, "club-houses") were established and provided local public goods. Yet, membership in each guild was limited to those from a specific place of origin. The *huiguan* thus extended the reach of the rural clans into the city. Guilds in the European sense (*Gongsuo* "Public hall"), in which membership was based on local residency and common economic interest, became common in China only in the 19th century (Christine Moll-Murata, 2008; Liu 1988). In fact we know of only 268 interest-based guilds that were created in China from 1655 to 1911 and most of them were founded in the 19th century (Christine Moll-Murata, 2008).

3.2.3   Institutions, exchange, and morality

The institutional foundations of markets in Europe and China also reflect distinct moralities. Market institutions differed due to the distinct social organization and morality. In Europe, intra-city formal enforcement institutions supported inter-city impersonal exchange through the Community Responsibility System (CRS). Under the CRS each and every member of a city was liable for a default by any other member who defaulted on his inter-city contractual obligation (Greif 2006b). Trade would cease following an uncompensated default motivating the city to compensate for the default and punish the one who defaulted to deter others. General morality renders this punishment, *ceteris paribus*, more credible than limited morality does.

Under limited morality a CRS is less effective. To see why, note that under general morality both intra-city and inter-city impersonal exchange requires to rely on legal contracting. This is not the case under limited morality, however. Exchange with a clan member – even from afar and without personal

34

familiarity – has the comparative advantage that limited morality entails in intra-clan exchange. It is reassuring to find that, in fact, long distance trade in China was conducted by clan-based and region-based commercial networks in which limited morality and informal institutions supported cooperation. "The sprawling merchant diaspores that managed 18th century interregional trade usually were built upon kinship ties. The huge shipments of rice to the lower Yangtze ... for example, were overseen by groups of Kiangsi merchants organized internally by lineage" (Rowe 2002, pp. 531-2; cf Ma 2004).[21]

Arguably, these distinct institutions and the implied patterns of interactions fostered their underpinning moralities. The Community Responsibility System implied equality before the law, thereby confirming this moral principal. Moreover, this system implied a shared interest in promoting generalized morality. A city's craftsmen and merchants were held by outsiders to be jointly liable for misconduct. Thus, they stood to benefit from socializing each other to generalized morality. Indeed, equal treatment of out-of-town traders is mandated in more than 50 percent of the available brokerage rules in merchant towns in Central and Western Europe from the late 13th to the end of the 17th century. This rule had a positive and significant impact on trade (Boerner and Quint 2010).

In contrast, limited morality in China was arguably reinforced by personal exchange in clan-based and region-based commercial networks. This positive feedback is reflected in clans' rule books from the early 20th century. Although "friendship is one of the five ethical relationships [in Confucianism] and should not be disregarded, yet [clans' rules from the 20th century often state that] one must be very careful about it" (Liu, 1959, p. 148). About 95 percent of clan rules call for care in selecting friends while only 8 percent call for "helping a friend in trouble" (Ibid).

---

[21]For Chinese trading networks abroad and the importance of intra-group morality in their functioning, see Landa (1981), Pyatt and Redding (2000), and Redding (1993), among others.

3.2.4 The scope of morality

Whether charity is more personal or impersonal is a good proxy for the scope of altruism and for whether obligations were felt towards members of the group or society as a whole. Charity is personal when the giver donates to specific individuals known to him or her and it is impersonal otherwise. If limited morality prevailed in China, personal charity among kin should have predominated. In contrast, if generalized morality prevailed in Europe, impersonal charity to non-kin should have predominated. This, indeed, was the case.

Prior to the 17th century, charity in China was predominantly personal, given to kin, and provided by the clan. The innovator of the clan trust (lineage estate), Fan Chung-yen (989-1052), "had ruled that the lineage should aid only relatives with lineage ties that were clearly documented in the genealogy" (Smith 1987, p. 316). There were some notable exception. Buddhist organizations dispensed some impersonal charity in China but their resources were limited after the persecutions in the 9th century. Similarly, the state-run granary system (that reduced price fluctuations) assisted the poor but did so only following a natural disaster that influenced many.

In China the first non-Buddhist organization dedicated to impersonal charity was established in 1590 and similar ones followed by the early 17th century. The Chinese authorities did not object impersonal charity organizations and actually encouraged the practice[22]. Objection was put forward by scholars who decried the immorality of diverting assistance way from kin. A popular 17th century morality book "tells of a generous scholar who was de-

---

[22]Indeed, inter-kin Buddhist charity organizations catered to animals prior to the 17th century. Smith (1987) summarizes: "during the late Ming and early Ch'ing (1580- 1750), the method and spirit of charitable giving changed in China.... The benevolent societies and foundling homes of the late Ming and early Ch'ing were locally sponsored, voluntary, widespread, and enduring. They should not be confused with the state-sponsored, sporadic, and temporary institutions of the Sung dynasty...[The new organizations were devoted] to voluntary or compassionate giving to the poor and needy outside one's family" (pp. 319-21).

rided by a member of his lineage for lightly giving money away to strangers"
(ibid). Only subsequently impersonal charity organizations "evolved into
lasting institutions with enormous buildings, large endowments of land, and
high visibility" (Ibid, p. 317).

In contrast, by the 16th century impersonal charity had already been long
established in Europe although it was undermined by the "increased interven-
tion of the state in the provision of help for the poor, and for all the changes
in the theology of "good works" that the Reformation entailed" (Krausman
2000, p. 320).[23] Yet, in 1560, there were about 35 poor relief-foundations in
London and their number increased to more than 100 by 1700.[24] Posthumous
inflation-adjusted bequests of large benefactors quadrupled in England from
1550 to 1650 (while the population only doubled). Donations for the "relief
of the poor" or the "poor man box" appear in a quarter of the wills in late
sixteenth century Bristol and a century later this fraction had been about 30
percent in England as a whole (Krausman 2000).

The low level of impersonal charity in China manifested itself also in
examples of extreme poverty and desperation that received no assistance
from society. The Ming dynasty recruited eunuchs to fill administrative and
military posts and poor men thus castrated themselves to be considered.
Although the court may have employed as many as 70,000 eunuchs (Fairbank
1994, p. 130), many applicants were declined (Robinson 2000). Those who
were rejected became marginal in a society concerned with maintaining the
blood line. Poor castrated kin were not high on the priority list for clan-based
charity. In the absence of impersonal charity, they resorted, as Robinson
(2000) has documented, to banditry.

Another phenomena revealing the lack of impersonal charity in China is
wife-selling; the sale of a wife by her husband, usually directly to another

---

[23]It goes without saying that personal charity was also practiced in Europe. Donations
to ransom captives, for example, were often provided to rescue particular individuals, often
a kin

[24]Population grew proportionally more.

man to become the latter's wife or concubine (Sommer 2007). Wife-selling was not due to wide-spread poverty as it predated the economic decline on the eve of the modern period. Although the practice was declared illegal during the Qing dynasty (1644-1912), it remained common. In the absence of impersonal charity, wife-selling was a desperate survival strategy; a husband sold his wife (often with their children) as a last resort in order to survive. Sommer (2007) analysis of 242 court cases from 1750 to 1900 reveals that poverty motivated more than 75 percent of the sales.

The morality of a pre-modern society can also be extrapolated from its criminal law because their such laws were based on the principle that the punishment should fit the crime. The European cities' law reveals generalized morality. Punishments did not depend on the relations among the parties involved. In contrast, punishments according to the Chinese law depended on relations and were more severe for crimes committed against one's kin. Generally, penalties were most severe for crimes committed against senior relatives, less severe if committed against those outside of the extended family, and least severe if committed by senior family members against their younger kin. Even in "the early twentieth century, a father could kill his son without incurring much more than a reproof and a warning, while a son who killed his father, or even his only slightly older brother, faced a very hard time. Only the killing of friend by friend came to court on an even keel, so to speak" (Gellhorn 1987, p. 2).

In contrast, European legal codes exhibits uniformity and emphasized equality before the law. In China, there was a large geographical diversity in intra-clan legal customs and codes that were specific to a village, clan, or a family group (Ruskola 2000, 164, fn 37). In European cities everyone was de jure equal under the positive (man-made) laws.[25] Moreover, everyone was under the 'normative' (divine) law. This is evident in European coronation

---

[25]Indeed, Code Napoleon (1804) particularly relied on the laws of Paris, Coutume de Paris. The comparable Civil Code of Austria (1811) was similarly based on the ideals of freedom and equality before the law.

ceremonies in which rulers were crowned by a representative of the Church to symbolize that they were under the normative law. In contrast, the Chinese law was explicitly unequal. Public officials and their kin had legal 'privileges.' An official, depending on his rank, and his kin, depending on their blood relations to him, were subject to a lesser penalty than other perpetrators for the same offense (McKnight 1985).

3.3   Society and State

Extending the historical discussion to incorporate the state transgresses the boundaries of our formal analysis. We nevertheless discuss this issue because it confirms the importance of the cultural and institutional bifurcation we focus on and highlights a political positive feedback that further contributed to it.[26]

The observation that pre-modern European rulers extensively harnessed cities' administrative capacity in pursuing their interests confirms their importance in sustaining cooperation. Rulers relied on cities to collect tax, provide navies, muster military contingencies, and administer justice. The implied administrative power of the European cities vis-à-vis the state and intra-city cooperation enabled most cities in Western Europe to gain self-governance by 1350 (e.g., Greif 2005).

The eventual transition in European states to more centralized administrative and legal structures was arguably facilitated by the generalized morality that the above system implied, its legal infrastructure and principles. In England, for example, the merchant law was integrated in the Common Law. In France, Code Napoleon drew heavily on the legal code of Paris. Similarly, the European self-governed cities became an integral part of states' administrative apparatus.

---

[26]Greif (2006) elaborates on this 'fundamental asymmetry' in institutional development. Institutional (including cultural) elements inherited for the past influence subsequent development due to such factors as shared cognition and expectations, organizational capacities, etc..

The importance of the clans in enabling cooperation in late Imperial China is similarly evident in their integration in the administration of the Chinese Empire. Although details varies, clans collected tax from their members, were legally liable for their criminal conduct, and cooperated with the magistrates in the provision of public goods. Reliance on clans implied that Chinese bureaucracy could withdraw "from official involvement in local affairs. ... Actually, the emperor's appointee to any magistracy could administer it only with the cooperation of the gentry [local elite] in that area" (Fairbank 1994, p. 106).

The policy of limiting the power of local magistrates was a response to rebellions by powerful governors that contributed to the fall of the Tang dynasty in 907. When the Sung dynasty reestablished central control in 960, the Emperors relied on a small administration.[27] As noted above, the number of public officials remained almost the same from the 10th to the 19th century despite a large increase in population and territory.[28] For such administration to function, it had to rely on local intermediaries, the clans and the elders and gentry that led them.

The Chinese legal system reflect the importance of the clan to which the state delegated legal authority. In China, "a localized lineage was normally the largest unit within which disputes between its members were settled" (Freedman 1966, p. 114). In other words, clans were expected to specify and enforce the rules governing the relations among their members.[29] In terms of our analysis, relying on this quasi-private enforcement was preferred by the central authorities despite the economics of scale in law enforcement

---

[27]The process of creating a centralized bureaucracy had began much earlier. E.g., Ma 2008; Kiser and Ka 2003.

[28]The number of counties increased from 1,230 under the Sung to 1,360 under the Qing while population increased from about 110 to 425 millions during that time. The total number of officials remained about 20,000. Fairbank (1994), p. 106. Officials relied, however, on unauthorized additional staff.

[29]Clans were not 'imperia in imperio' and dissatisfied clan members could appeal, and did appeal, to the Imperial legal system (Ruskola 2000, p. 1663).

because clan enforcement took advantage of the disciplinary impact of limited morality and reputation.

The well known consequence is that the Chinese state invested relatively little in legal infrastructure and its executives (the magistrates) were also responsible for adjudication. Our analysis highlights why. An effective and impartial legal system would have undermined the clans, an outcome contradictory to the interests of the elders who controlled the clans and the state that used them. Thus, recognizing the administrative role of intra-clan enforcement explains why China did not develop commercial legal code until the late 19th century and why civil legal adjudication was aimed at finding a compromise.[30]

Indeed, the Chinese state encouraged intra-clan disputes resolution while clan rules regularly discouraged litigation and favored arbitration provided by the clan (Liu 1959). In terms of our analysis, limited morality reduced intra-clan enforcement cost and undermined the potential gains from state-provided legal institutions. Limited morality and a tradition of private adjudication and unequal legal rights, however, have arguably reduced the gain from creating a civic legal system. The large extent to which the law was delegated is evident from the observation that in 1923, most of the land and buildings in Canton only had private contracts, namely, contracts signed by seller and buyer, without 'red deeds,' that is, sealed Imperial deeds (Chung 2010. p. 1431).

Because Chinese clans economized on enforcement cost, the state created complementary institutions to pursue its objectives. It promoted the ideology of filial piety, the moral obligation toward patents and ancestors. Although filial piety had been a cornerstone of Chinese ethics since the Han dynasty (206 BCE– 220 CE), but the Sung developed, codified, and popularized it (Hamilton 1984). More generally, Late Imperial China promoted

---

[30]The state enforced, however, legal rights over taxable land. For illuminating discussions of the legal system and the scholarly debate about its nature see also Zelin (2004), Nakamura (2004) and Ma (2007).

Neo-Confucianism. in which "the family was given a metaphysical foundation, and filial piety was promoted" (Ruskola 2000, p. 1622).

The Chinese state reinforced intra-clan cohesion by regulations such as linking rights to buying land to local clans' membership, regulating geographical mobility, and rendering the lineage collectively responsible for crimes committed by a member. In general, laws backed the power of the kinship organizations – and the elders who controlled them – over their members.

4   Differences within Europe and China

This section exploits heterogeneity within Europe and China and presents additional historical and contemporary evidence that urbanization is less diffuse in areas where kin-based relations are stronger and better organized.

4.1   Urbanization in the European regions

As a measure of the strength of kin-based relations, we rely on the family structure. Todd (1990a, b) documents that, during the Middle Ages and earlier, different parts of Europe developed distinct traditions of family structures, that remain relevant today. Some of these family structures share similarities with the clan, others do not. In the extended family, several generations live together under the same roof, respect patriarchal authority, and engage in risk sharing and other forms of cooperation, pretty much as in the Chinese clan although on a smaller scale. In the nuclear family, by contrast, adult children are emancipated and leave the family home, and cooperation and risk sharing within the family, while still present, are less pronounced.

We take these different family traditions as proxying for the distinction between kin-based vs generalized moral obligations. Under this assumption, our theory predicts that, in regions where the extended family prevailed, individuals were less easily attracted to the city, compared to regions where instead the nuclear family was more diffuse. To test this prediction, we

collected data on urbanization rates of different European regions around 1600. We then ask whether these historical urbanization rates are explained by the diffusion of different family types in the region, controlling for other observable regional features.

The theory also predicts that urbanization in turn favors the diffusion of the nuclear family and a gradual erosion of the extended family tradition. This is why we go back in time to the period before the industrial revolution, when only a small fraction of the European population was urbanized: in our sample on average only 8-9 percent of the population lived in cities of at least 10 000 inhabitants around 1600. Moreover, Todd argues that his classification of family structures captures remarkably stable traditions dating to the Middle ages if not earlier. Nevertheless, the possibility of reverse causation cannot be ruled out, so the regressions below can be interpreted as correlations consistent with the theory, not as strictly causal relations going from family structure to urbanization.

4.1.1 Data

Our data is built starting from Duranton and Rodriguez-Pose (2008), who in turn draw on Todd (1990). The sample includes 11 countries: Austria, West Germany (the former Federal Republic of Germany), Spain, Finland, France, Ireland, Italy, the Netherlands, Portugal, Sweden and the United Kingdom. Depending on the variables used, the sample ranges from 70 to 83 regions, although we report results for a core sample of 74 regions.

Following Duranton and Rodriguez-Pose (2008), regions are defined at the NUTS II level, according to the 2003 Eurostat classification. However, in order to keep data consistency when matching different sources, we have aggregated some regions into larger units - see below and the data appendix for more details.

Urbanization around 1600 The dependent variable is the percentage of regional residents that, around the year 1600, lived in cities of at least 10.000 inhabitants (*urbanization around 1600*). The numerator, city size in 1600, comes from Bairoch, Batou and Chévre (1988). The denominator, regional population around 1600, is constructed from a variety of sources. For Spain, the UK, Sweden and the Netherlands we have precise estimates of regional population around 1600, from specific sources listed in the appendix. For the remaining countries (Austria, Finland, France, Ireland, Italy, Portugal and West Germany), the earliest regional population data we could find dates to 1860. We also have estimates of national population in 1600, from McEvedy and Jones (1978).[31] We have thus imputed population in 1600 to each region of these countries under the assumption that the quota of national population living in each region has remained constant between 1600 and 1860. In the four countries for which regional populations can be obtained in both ways, the correlation between the imputed and the historical data on regional population is 85 percent. This suggests that the assumption of a constant partition of national populations across regions is not too restrictive.

The distribution of *urbanization around 1600* in our sample is illustrated in Figure 5. It has a mean of 8.8 percent and it ranges from 0 (for several regions) to 51 percent, for Andalucia.

<<Insert Figure 5 about here>>

Family types Following Todd and Duranton and Rodriguez-Pose, our classification of family types centers on the key distinction between the extended vs the nuclear family. These two family structures differ in the degree of cooperation between subsequent generations, and in the authority exercised by parents. At one extreme, nuclear families are those in which children are

---

[31]Their classification of nations coincides with the current repartition of countries, with some differences leading to the definition of the nations of Belgium-Luxembourg, Republic of Ireland-Northern Ireland, England-Wales, and Scotland.

emancipated from their parents and leave the household at the time of marriage or before. Each family thus consists of at most two generations. At the opposite extreme, the extended family typically consists of three generations living together and mutually cooperating under patriarchal authority.

Todd (1990a) measures the diffusion of both family types across Western Europe. His starting point is the census data of 1960-1980, from which he constructs an index of the number of households composed of at least three cohabiting generations (grandparents, married parents, children). Having removed the effect of rural employment, he identifies areas where the extended family is comparatively more frequent. He then combines this contemporaneous information with a large number of historical studies and qualitative information going back several centuries, to obtain a final classification of areas within Europe where the extended or nuclear family traditions prevailed throughout history (in some areas neither family type was dominant, or no clear classification is feasible). These areas reveal considerable heterogeneity within countries, and often country borders do not overlap with the areas where one or the other family tradition is dominant.[32]

Duranton and Rodriguez-Pose (2008) map Todd's original classification into the current administrative units corresponding to the European regions, and their work is what we rely upon. Specifically, the variable *extended family* is the fraction of area in the region where the extended family prevails. Duranton and Rodriguez-Pose (2008) start out with NUTS III regions, and aggregate them up into the larger NUTS II regions by weighting observations with the land area of each region. As explained above, in some cases we impose an even higher level of aggregation, again weighting observations by land area to maintain consistency with Duranton and Rodriguez-Pose (2008).

Some areas are classified by Todd (1990a) and by Duranton and Rodriguez-Pose (2008) as being dominated by no particular type of family, or as not

---

[32]Todd (1990a)'s classification is actually finer than the one we use, since he also classifies families according to whether they treat children equally or unequally. Since this distinction is not relevant for our purposes, we neglect it.

classifiable, or for which information is incomplete. In our core samples, we have excluded the 9 regions in which more than 50 percent of the area is either unclassified or where no family type is dominant. The results are robust to including these 9 regions in the sample, and to excluding another 4 regions where a positive but minoritarian fraction of the area is either unclassified or where no family type is dominant.

The distribution of the variable *extended family* in our core sample is illustrated in Figure 6. It has a mean of 60.7 percent, and ranges from 0 to 100 percent.

<<Insert Figure 6 about here>>

Contrasting Figures 5 and 6, regions where the extended family prevails (darker regions in Figure 6) seem to be associated with a lower rate of urbanization (lighter regions in Figure 5), as predicted by the theory.

Other variables   We also collected several historical variables that may influence regional urbanization. Specifically, the variable *Sea* is a dummy variable identifying regions bordering with the sea. The *Atlantic Ocean* dummy identifies regions touched by the Atlantic Ocean, the Celtic Sea or the North Sea, except for Swedish regions. *Longitude* refers to the longitude of a point within the region, corresponding to the capital, the largest city, or an average of the capitals of the districts within the region. *Population density* refers to regional population around 1600 (as defined above) divided by the area of the region expressed in square kilometers. The dummy variable *Capital* identifies regions containing the current capital city of the country, with Bonn as the capital of West Germany.

Since urbanization rates are likely to be explained by several unobserved variables that may vary at the national level, all regressions also include country fixed effects, with countries defined according to the situation around 1600. Specifically, southern and insular regions of Italy have been coded as

46

belonging to Spain, and the NUTS I level region of East France (consisting of Alsace, Lorraine and Franche-Comté) is classified as belonging to Germany.

4.1.2 Results

Table 3 presents the results. The dependent variable is always *urbanization around 1600,* country fixed effects are always included. Standard errors are clustered by country (with the same definition of country as used in the fixed effects), to allow for measurement error or omitted variables correlated across regions belonging to the same country.

Column 1 contains the more parsimonious specification, and additional regressors are added in each column. The estimated coefficient of the variable of interest, *extended family,* is quite stable and always statistically significant at the 5 percent level (significance increases to 1 percent if the standard errors are not clustered). As predicted, diffusion of the extended family is associated with a reduction in the rate of urbanization. The quantitative effect is also relevant: if this was a causal effect, as *extended family* switches from 0 to 1, *urbanization around 1600* increases by 11 percentage points, corresponding to more than doubling the sample mean, and almost one standard deviation of the dependent variable.

<<Insert Table 3 around here>>

Figure 7 illustrates the correlation in the data (after removing the effect of all regressors included in the least parsimonious specification of column 4 in Table 3). Several outlier observations are present, but none of them individually is driving the correlation.

<<Insert Figure 7 around here>>

Finally, the results are robust to alternative specifications (not shown), such as replacing the variable *population density* with the land area of the region (or adding land area to the specification in column 4), expressed in

square kilometers or in logs. The results are also robust to alternative definitions of urbanization (such as changing the city size behind our measure of urbanization from 10.000 inhabitants to 15.000 or 5.000, or changing the reference year in which it is computed, or changing some of the least reliable data sources).

Altogether, these results confirm the correlations predicted by the theory: in European regions where family traditions denote stronger kin-based obligations, the rate of urbanization was slower and more sparse.

4.2 Clan Persistence and Urbanization in Modern China

On the eve of the modern period, as noted above, clans were still prominent in China. The modernization movement in late 19th century, however, viewed clans an obstacle to economic development and the communists officially abolished them after gaining power in 1949. Clans' properties were confiscated, elders lost their legal privileges and authority, clan legal codes were no longer recognized, and the ideology of class consciousness was promoted (e.g., Haung 1985, p. 308).

Had the clan been a social organization without a moral component, this policy would have put an end to it. Yet, the clan survived as our theory predicts. This section presents supporting evidence from recent censuses and establishes that in contemporary China areas with stronger clans are less urbanized. Specifically, there is a statistically and economically negative correlation between county-level urbanization rate and proxies of clannishness. Both clans and urbanization rate are endogenous in our model and reverse causation is thus possible.[33] The evidence should be interpreted as correlations consistent with the theory, and not as causal relations.

_____

[33]This specification captures that migration is an action an individual can take while the existence of a clan organization is beyond the control of each individual member.

4.2.1 Data and variables

Our data is taken from two sources, a random sample of 76 counties from *China County-prefectural Statistical Yearbook*, 2006 (YB06) and *China General Social Survey*, 2005 (GSS05), a corresponding geographically-based random sample of individuals.[34] Although the sample is small, the data set contains information on 4274 individuals from 205 villages. Counties are, roughly speaking, the third administrative units in China and as such contain both townships and villages but not cities that are on higher, provincial or prefectural, levels. Our analysis thus pertains only to intra-county migration to townships and not cities. Internal migration has traditionally been limited by the Household Registration System, but enforcement has been gradually relaxed since the 1990s and in our sample about 6 percent of the individuals are immigrants.

The dependent variable (*urbanization)* is the 2006 county-level share of the population living in urban communities. An area is urban, by the year book definition, if, roughly speaking, its population density is at least 1,500 persons per sq.km or it is the seat of the local government.

The main independent variable is a county-level measure of clan-based organizations (*clan organization*). The GSS05 asks (only) rural residents whether there is a clan network or organization in their community and, if there is one, if it is a surname-based or a temple-based one.[35] (Henceforth surname-based (or informal) and temple-based (or formal) organizations respectively.) The data reveals that the two types of clan organization are almost identical in terms of their ritual-related assets. More than 90% of organizations of either type have a genealogy, a graveyard or both. The two

---

[34]See http://www.ust.hk/~websosc/survey/GSS_e.html
or http://search1.odesi.ca/details/view.html?q=social+networks&field=KW&uri=/icpsr/23861.xml.
and Chinadataonline.com.

[35]F10. Is there any clan network or organization in your community/village? 1 None. 2 There are kinship network bounded by family name, but no formal organization. 3 There are clan organization centered around clan temple activities.

types differ, however, in their economically-relevant assets such as land, estates (other than ancestral hall), and trust funds. Only 26% of the surname-based network have such an asset compared to 78% of the temple-based organizations (qf12).

The data allows for multiple proxies of clannishness. Although underreporting of clan organizations is likely given tradition of suppression by the communist authorities, the average number of clan organization per-village is 1.35 and 40 percent of villages have at least one clan organization. Given that an average village has 30,000 inhabitants and 20 respondents, the multiplicity of reported clan organization per-village can be due to either the existence of multiple organizations or multiple reports of the same organization. Because we cannot rule out that the same organization is reported by more than once, our main proxy of clannishness is the county-level probability of having at least one clan organization in a village.[36] In constructing this variable, a village with at least one (formal or informal) clan organization gets a value of one (and zero otherwise). The county-level clan organization variable is the mean of these village-level dummy. This proxy is far from perfect. The observation that individuals from the same village differ on whether there is a clan organization in their community suggests the limit of the village as a unit of analysis. At the same time, a strong and negative correlation between clan organizations and women political participation (in the CGSS) suggests that the presence of clan organization captures a village-wide cultural feature (Kung and Ma 2010). In any case, we check the robustness of our analysis using various alternative proxies (see appendix). Among them is an individual-level proxy which is the probability of membership in a clan organization. It is the county average of the per-village proportion of the respondents who reported a clan organization.

In our model, migration is also influenced by characteristics of the individ-

---

[36]A village has a clan organization if at least one individual in that village said that there is such an organization.

ual and the city. Accordingly, we use the following controls: the county-level, village average years of education (*education*), village average rural household income (*income*), total investment in fixed assets in urban areas (*urban fixed investment*), population density (*population density*) and dummies for China's six regions (*region*).[37]

The following table provides summary statistics and the data appendix provides additional details.

<<Insert table 4 here>>

4.2.2 Results

In our 205 villages, 277 individuals reported clan organizations implying, on average, one organization per 15.5 respondent and one formal organization per 53 people. The highest number of people per-clan organization is in the eastern region (35) and the lowest is the northwest (8). There is at least one clan organization exists in 40% of the villages and a formal organization in 13.6% of them. On the county level, 65% of the counties have clan organization(s) and in 41% of them the probability of having one is at least 50%. Almost 70% (2961/4724) of the population live in a county in which the sample probability of having an organization is positive. Clearly, the clan organization is still a feature of China's social fabric.

Consistent with the historical evidence that clans prevailed throughout China, clan organizations currently exist in each of China's six regions although there are no formal organizations in the north (figure 8). This last observation is in line with the above historical analysis. Clans were initially formed as people migrated from the north to the south and west. Similarly, formal organizations are particularly strong in the south central region and, specifically, in Guandong,the most populated and richest province.

---

[37]Pre-moden China expanded inland and not by sea. In any case, the analysis is robust to a dummy for coastal provinces. The difference between rural and urban income is also insignificant.

<<insert figure 8 around here >>

Our preliminary regression analysis (not reported) reveals that there is no significant relations between informal (surname-based) organizations and urbanization (at the 10% level) but urbanization rate and formal (temple-based) organizations are significantly and negatively correlated. Specifically, the estimated coefficient is stable, statistically significant at the 1 percent level, and does not decline when additional controls are introduced while the R=squared more than triple, to reach 37 percent. The impact of clan organizations is also economically meaningful. After controlling for the other variables (model 5 in table 5), an increase of one standard deviation in clan organization is associated with a 5.6 % reduction of urbanization rate and a ten percent increase from the mean is associated with a 1.1 % reduction of urbanization rate.

Table 5 presents the OLS regression results while omitting one county that our preliminary regression analysis identified as an outlier in favor of our hypothesis.[38] Column 1 contains the more parsimonious specification of clan organizations and education. Additional regressors are added in each column. The last column controls for regional fix effects.

<<Insert table 5 here >>
<< Insert figures 9 >>

Figure 9 illustrates the partial correlation (meaning, after removing the effect of other regressors (model 5). Several outlier observations are present, but none of them individually is driving the correlation. As noted above, our model highlights the difficulty of attributing causality. Yet, the finding that only formal organizations – that have economic assets – matters raises the question of whether the correlation is only due to the economic benefits of these local-specific assets. The evidence, however, suggests that this in

---

[38]County 452723 in Guangxi province. Its predicted residual is more than twice the next largest residual.

not the case. Controlling for age, education, and gender the income of those who reported a formal clan is not statistically different from the population at large.

The results reported in Table 5 are robust to alternative specifications (not shown) such as using land area instead of *population density,* gdp_pc instead of rural-household income, proximity to the coast instead of regions and excluding non-Han or non-local people.[39] The result is also not due to the impact of the south central region in which, as noted above, clans are particularly prevalent. Estimated separately, there is a significant and negative correlations between clannishness and urbanization in the (22) counties in that region and in the (53) counties in the rest of China. These correlations also hold when the proxy for clannishness is the probability that an individual is a member of (i.e., report knowing of) a formal clan organization. These results are presented in the appendix.

Altogether, these results illustrate the persistence of the clan and confirm that where clans are stronger the rate of urbanization is lower.

5 Clans, Cooperation, and Institutions in Post-Reform China

The reemergence of clan organizations and other relations-based organizations in post reform China further confirms complementarity between limited morality and clannishness. Clans and clan-like organizations re-established themselves as the locus of cooperation and an integral part of various institutions.[40]

Although after 1949, the Chinese state promoted the ideology of class

---

[39]Specifically, the result (model 5) hold by constructing the clanorg measure based on the sub-village units rather than the villages, when a county index gets the value of its village-highest clanorg value, and when the village-level index is construted only for villages in which at least two individuals reported a clan organization.

[40]See Greif and Tadelis (2010) for a theory of moral persistence.

consciousness and class-based morality, clan- and relations-based morality perpetuated. Limited morality was actually strengthened by some policies such as legal restrictions on geographical mobility, collectivization, and the delineation of village boundaries (e.g., Haung 1985). In terms of our model, these policies fostered limited morality by forcing kin to remain attached to the clan. Kinship remained important. To illustrate, in the 1980s the state still employed four-fifths of the labor force and job assignment was formally based on meritocracy. Yet, in a representative sample of 1,008 households in the city of Tianjin (1988), 57% reported that they got their jobs (1977–1988) based on personal connections (*guanxi*) and 43.3% of those who assisted them were relatives (Bian 1994).

More recent pieces of evidence similarly reveal limited morality. The GSS06 directly asked villagers how much they trust an individual with the same surname vs a villager with another surname (qf14). An unconditional mean comparison test of the resulting 4228 responses reveals a statistically significant lower trust (at the 1% level). More generally, less than 40 percent of contemporary Chinese trust strangers while more than 60 percent do in the US (WVS, Roland Inglehart, et al 1998).[41] Only 11.3 percent of Chinese trust a person whom they met for the first time compared to between 26.1 percent to 49.3 percent in the West (i.e., France, GB, USA and Germany).[42] While 90 percent of Chinese do not trust (much or at all) a person whom they met for the first time, only 60 percent feel the same in the West. Trust in people that one knows personally is only 20.5 percent in China, but 31.6 percent in the West excluding France (and 38.8 percent including it). Trust in neighbors is high both in China and the West, but China stands out in having 27 percent who 'trust completely' their neighbors compared to an average of 14.4 percent in the West excluding France (and 20.1 percent including it).

---

[41] Herrmann-Pillath 2009 discusses the difficulty of international comparison of trust. His insightful analysis of the evidence, however, is consistent with our analysis.

[42] This paragraph is based on the World Value Survey, 2005-8. www.worldvaluessurvey.org.

While friendship is 'very important' to less than 30 percent of Chinese, it is very important, to almost 60 percent in the West.

If social organization and morality are complements, as we claim, clans should have been resurrected following the market reforms that allowed individuals to organize themselves. This, in fact, has been the case. A county-level survey (2000) in the southern province of Jiangaxi, for example, reveals that 70 surnames out of 99 (in 40 villages) updated their genealogies since 1981. Moreover, from 1991 to 2000, 41 surnames repaired or rebuilt their ancestral shrines (Liangqun and Murphy 2006, p. 630). Of the 887 households that moved to or from one of the 50 villages in that county, more than 90 percent relocated to their 'ancestral village' and 60 percent relocated due to inter-lineage tension (ibid, p. 623). A representative national survey of 316 villages reveals lineage activities and kinship organizations in 66 percent of the villages (from 1999 to 2002). Since the reforms, 19 percent of these villages have constructed or refurbished ancestors' halls (Tsai 2007a, pp. 154–7). It is not uncommon that collectively owned rural firms formally exclude non-locals (Thøgersen 2002).

The reemergence of clans is particularly noteworthy given that the reforms were not designed to foster clan organizations.[43] Households, and not clans, were given land user rights in the former collective farms and privately-owned businesses were permitted. Yet, kin-based and relations-based exclusive organizations have reemerged and resumed their traditional role in supporting cooperation. Table 6 presents the number and fraction of organizations – out of the 277 in our sample – that fulfil various functions. Most common are cooperation-promoting functions– resolving private disputes within the village and handling inter-village relations – and providing public goods in the village or the clan.

<< insert table 6 around here>>

---

[43]Yet some policies under communism actually strengthened limited morality. Among these policies were legal restrictions on geographical mobility, collectivization, and the delineation of village boundaries (e.g., Haung 1985).

Previous studies have already noted the importance and roles of clans in securing property rights from local officials, organizing weddings and funerals, providing welfare, contributing to public projects (e.g., road construction), and promoting mutual aid arrangements. For example, in River county (Jiangaxi) descent groups re-grouped to protect land-user rights (Liangqun and Murphy 2006, p. 632) and village-wide lineage groups have a positive, large, and statistically significant impact on the provision of public goods and they hold public officials accountable (Tsai 207). There is a strong and significant effect of village-level kinship on the number of private enterprises and the labor force. A 10 percent increase in the proportion of households belonging to the largest lineage group implies a 33 percent increase in the number of private enterprises (not including self-employed individuals) and a 24 percent increase in their average workforce.[44]

More generally, the reform led to proliferation of institutions based on limited, relations-specific morality and affinity such as *guanxi* (relations) and clan-like organizations.[45] In contemporary China "you trust your family absolutely, your friends and acquaintances to the degree that mutual dependence has been established ... With everybody else you make no assumptions about their good will" (Redding 1993, p. 66).[46] Inter-clan cooperation is commonly based on *guanxi* (relations; e.g., Thøgersen 2002, p.

---

[44]Peng (2004) based on survey of 366 villages from 22 counties (1993-4). In the sample, 22% of households belong to the largest lineage group and 40% belong to the top three groups. Including self-employed the corresponding effect is smaller but significant; 7 percent increase per 10 percent increase in the proportion of the largest lineage group.

[45]In Late Imperial China, business were often formed based on fictive clan whose members were expected to behave as if they were related by blood (Faur 1989 and Rowe 2002, pp.355-7). Inter-clan business associations thus remained based on personal relations (Hamilton 1984).

[46]Similarly, outside of China proper, Chinese businessmen perceive Westerners as more reliable in contractual obligations. For example, a 1994-5 survey of Chinese businessmen in Thailand and Hong-Kong finds that "Westerners are considered [by the Chinese] to be attractive partners for ... their respect for the law and keeping of promises. Trustworthiness is a frequently mentioned trading attribute of non-Asians" (Pyatt and Redding 2000, p. 59) and the Hong-Kong Chinese businessmen consider the Thai-Chinese more opportunistic and no more trustworthy than Westerners.

261). *Guanxi*, however, reflects and reinforces limited morality. "To make such networks operate reliably, Chinese society has come to attach central important to the notion of trust. What is Chinese about this trust, however, is that it is ... limited to the partners in the bond. It works on the basis of personal obligations, the maintenance of reputation and face, and not on any assumption that a society's shared faith makes all who share it equally righteous regardless of whether you know them or not" (Redding 1993, p. 67).

6   Concluding Discussion

Positive feedbacks amplified distinct initial social and cultural conditions that prevailed in Europe and China a millennium ago. The resulting organization of society influenced subsequent social, moral, and institutional developments. In China, clans became the locus of cooperation among kin motivated by limited morality and informal institutions. In Europe, cities became the locus of cooperation among non-kin motivated by generalized morality and formal institutions. Clans and cities as the basic units of cooperation influenced the institutional foundations of markets, the provision of social safety nets, legal development, and the organization of the state. The institutional embeddedness of these social groups further reinforced their organizational forms and the moralities in multiple ways that our parsimonious model does not capture. These effects persist today.

Persistence, however, does not mean stasis and both system were successful in adapting to change. Yet their dynamics differed due to the complementarity between their cultural, social and institutional elements. Each system responded to changing needs by reshaping and recombining these elements.[47] For example, the Europeans responded to the technologically driven increase

---

[47]For this view of culture see, for example, Greif 1994, 2006 and Herrmann-Pillath 2010. For a discussion of the Chinese case, see Weber 2007; Thøgersen 2002; Herrmann-Pillath 2009; Redding and Witt 2009; Redding and Witt 2009.

in economies of scale by relying on the law and general morality to separate between ownership and control while in China, kinship organizations were supplemented by "more flexible, selective, and contingent networks that include other kin and relatives" (Thøgersen 2002, pp. 267).

Thus, the European system has a comparative advantage in supporting impersonal exchange while the Chinese system comparative advantage is in economic activities in which personal relations are more efficient. The relative efficiency of the Chinese and European systems is thus theoretically unclear. More generally, Europe and China recent economic, social and political histories bear the hallmark of their distinct social organizations and cultural heritages. In particular, Europe's transition to a modern economy and its successful post World War II reconstruction benefitted from its social welfare policy routed in general morality and formal institutions. The subsequent dynamic of the welfare state led to entitlement programs that, in some European states, are fiscally unsustainable. In contrast, post-reform China experienced rapid economic growth despite weak formal institutions, from a Western perspective, because limited morality fosters individual-level profit-seeking behavior on the one hand and relations-based, informal institutions for collective action and enforcement on the other.[48] Yet, limited morality and informal institutions foster potentially destructive inequality, corruption, and social tensions.[49]

More generally, our analysis illustrates the importance of social groups who organize themselves into basic units of cooperation. The impact of social groups on institutions is larger than the impact of their individual members because of the comparative advantage that intra-group morality entails in facilitating cooperation. Multiple institutions – internal and external to these groups – were created or emerged to complement or use this comparative

[48] E.g., Yu and Zhang (2008), Whyte (1995, 1996), Nee (2010).

[49] E.g., Nee (1992), Lau, Yingyi and Roland (1999), Zhang (2008), Li, Grace (2009), Su and He (2010), Li, et al (2008), Lu et al (2010), Thøgersen (2002), Bai et al, (2006), Lu and Tao (2010),

advantage. Different social groups can emerge and perpetuate due to positive feedbacks between social organization, morality, and institutions. At the same time, social groups are endogenous and their membership, morality, and impact depends on environmental factors that influence their members' incentives to take the actions required for their perpetuation.

The economic literature tends to group individuals based on their economic interests but the importance of social groups – and the moral obligations their members' share – is transparent in, for example, inter-group riots, inter-racial crime rates, intra-group economic cooperation, and patterns of charity giving. Attention to the role of social groups in the organization of society and economic and social outcomes is likely to lead to a better understanding of the outcomes we seek to comprehend.

7 Appendix

7.0.3 Proof of Lemma 1

Exploiting the results and the notation in the text, we have:

$$
\begin{aligned}
W^{\lambda n}(x^{\lambda n}, x^{\gamma n}) &= 1 + \lambda - \tau + H(g^n) \\
W^{\gamma n}(x^{\lambda n}, x^{\gamma n}) &= 1 + H(g^n) \\
W^{\lambda y}(x^{\lambda y}, x^{\gamma y}) &= \delta + 1 - q + H(g^y) \text{ if } x^{\gamma y} \geq \hat{e} \\
W^{\gamma y}(x^{\lambda y}, x^{\gamma y}) &= \delta + 1 + \gamma - \tau + H(g^y) \text{ if } x^{\gamma y} \geq \hat{e} \\
W^{\lambda y}(x^{\lambda y}, x^{\gamma y}) &= W^{\gamma y}(x^{\lambda y}, x^{\gamma y}) = \delta + 1 \text{ if } x^{\gamma y} < \hat{e},
\end{aligned}
\tag{9}
$$

where g^n and g^y are known functions of x^{pz} through (1)-(2). As stated in the text, δ^p is obtained from (9) setting $W^{pn} = W^{py}$. Exploiting (9), some simple algebra completes the proof.

7.0.4 Proof of Proposition 2

We consider each equilibrium separately

Equilibrium with Full Sorting By (9), the value of δ that leaves type γ indifferent between the clan and the city is:

$$
\begin{aligned}
\delta^\gamma &= \tau - \gamma + H(g^n) - H(g^y) \\
&= \tau - \gamma + H(\tau x^\lambda) - H[M\tau(1 - x^\lambda) - e]
\end{aligned}
\tag{10}
$$

for $x^\gamma = 1 - x^\lambda \geq \hat{e} \equiv e/M\tau$ where the second equation follows from (1), (2), (6). Clearly, equation (10) defines an implicit function $\delta^\gamma = D^s(x^\lambda)$, where the s superscript is a reminder that this is the full sorting equilibrium. By (10)

$$
\frac{\partial \delta^\gamma}{\partial x^\lambda} \equiv D_x^s = \tau[H_g(g^n) + M H_g(g^y)] > 0
\tag{11}
$$

The equilibrium conditions discussed above require:

$$
0 \geq D^s(x^\lambda) \geq 1/d - a
\tag{12}
$$

which by (11) is satisfied only for some values of x^λ. Specifically, implicitly define $\bar{x}$ and $\underline{x}$ by:

$$
\begin{aligned}
D^s(\bar{x}) &= 0 = \tau - \gamma + H(\tau\bar{x}) - H[M\tau(1 - \bar{x}) - e] \\
D^s(\underline{x}) &= 1/d - a = \tau - \gamma + H(\tau\underline{x}) - H[M\tau(1 - \underline{x}) - e]
\end{aligned}
\tag{13}
$$

By (11), $\bar{x} > \underline{x}$. Then an equilibrium with full sorting exists if $x^\lambda \in [\underline{x}, \bar{x}]$. Furthermore, (10) implies that

$$
1 - \hat{e} > \bar{x} > 1/2 > \underline{x} > 0
$$

if the following conditions are satisfied (each condition corresponds to one of

the above inequalities in the same order):

$$\tau - \gamma + H(\tau(1 - \hat{e})) \; > \; 0 \tag{A1}$$

$$\tau - \gamma + H(\tau/2) - H(M\tau/2 - e) \; < \; 0 \tag{A2}$$

$$\lambda + q - \tau + H(\tau/2) - H(M\tau/2 - e) \; > \; 1/d \tag{A3}$$

$$\lambda + q - \tau - H(M\tau - e) \; < \; 1/d \tag{A4}$$

Finally, note that in this equilibrium with full sorting, the fraction of each dynasty living in the city is

$$x^y = x^\gamma = (1 - x^\lambda) \equiv G^s(x^\lambda) \tag{14}$$

hence it moves one for one in the opposite direction with changes in the fraction of clannish types in the population.

Summarizing the above discussion:

Suppose that (A1)-(A4) hold. Then an equilibrium with full sorting exists if and only if $x^\lambda \in [\underline{x}, \bar{x}]$, where $1 - \hat{e} > \bar{x} > 1/2 > \underline{x} > 0$. In this equilibrium the fraction of each dynasty living in the city is $x^y = G^s(x^\lambda)$ given by (14), and it varies inversely with x^λ over the range $x^y \in [(1 - \bar{x}), (1 - \underline{x})]$

Equilibrium with segregation in the clan Next, consider the equilibrium where the clannish types are segregated in the clan, while the generalists are present in both communities. Repeating the previous steps, the value of δ that leaves type γ indifferent between the clan and the city is:

$$\begin{aligned}
\delta^\gamma &= \tau - \gamma + H(g^n) - H(g^y) \tag{15} \\
&= \tau - \gamma + H(\tau x^\lambda) - H[M\tau(1 - d\delta^\gamma)(1 - x^\lambda) - e]
\end{aligned}$$

61

where the second equation follows from (1), (2), (6), having used

$$x^{\gamma y} = x^\gamma - x^{\gamma n} = (1 - d\delta^\gamma)(1 - x^\lambda)) \tag{16}$$

by (6). Again, equation (15) defines a known function $\delta^\gamma = D^c(x^\lambda)$. By the implicit function theorem applied to (15):

$$\frac{\partial \delta^\gamma}{\partial x^\lambda} \equiv D_x^c = \frac{\tau[H_g(g^n) + M(1 - d\delta^\gamma)H_g(g^y)]}{1 - d(1 - x^\lambda)M\tau H_g(g^y)} > 0 \tag{17}$$

where the last inequality follows from (C1) and from the fact that in this equilibrium $\delta^\gamma \leq 1/d$. The equilibrium conditions discussed above require that the γ types are present in both the clan and the city. This requires

$$1/d > D^c(x^\lambda) > 0$$

In fact, the equilibrium conditions are more stringent than that, because, city size cannot be smaller than $e/M\tau \equiv \hat{e}$ in order to sustain the enforcement technology. Imposing this additional constraint, we have that equilibrium requires that $x^\lambda < 1 - \hat{e}$ and, by (16), that:

$$(1 - \frac{\hat{e}}{1 - x^\lambda})/d > D^c(x^\lambda) > 0 \tag{18}$$

Since $a > 1/d$, (18) then also implies that the clannish types are in all in the clan ($\delta^\lambda > 1/d$).

Note that, by (15) and (13), at the point $x^\lambda = \bar{x}$ we have $D^c(\bar{x}) = D^s(\bar{x}) = 0$. By (17), then, it follows that condition (18) is satisfied for at least some $x^\lambda \geq \bar{x}$. Since under (A1) we have that $1 - \hat{e} > \bar{x}$, we know that an equilibrium with segregation in the clan exists for at least some $1 - \hat{e} > x^\lambda > \bar{x}$. Let $x^{\max}$ be the upper bound for x^λ where (18) is satisfied. By (15) and by the definition of $\hat{e}$, this upper bound is implicitly defined by

the condition:

$$(1 - \frac{\hat{e}}{1 - x^{\mathrm{max}}})/d = \tau - \gamma + H(\tau x^{\mathrm{max}}) \tag{19}$$

Clearly, $x^{\mathrm{max}} < 1 - \hat{e}$. Moreover, since at the point $x^{\lambda} = \bar{x}$ we have $D^c(\bar{x}) = D^s(\bar{x}) = 0$, since $D_x^c > 0$, and since by (19) $D^c(x^{\mathrm{max}}) > 0$, we must also have $x^{\mathrm{max}} > \bar{x}$. Hence if $x^{\lambda} \in (\bar{x}, x^{\mathrm{max}})$ this equilibrium with segregation in the clan exists.

Finally, note that in this equilibrium, the fraction of each dynasty living in the city size is

$$x^{\gamma y} = (1 - dD^c(x^{\lambda}))(1 - x^{\lambda})) \equiv G^c(x^{\lambda}) \tag{20}$$

Differentiating with respect to x^{λ} we immediately have that, by (17),

$$G_x^c = -(1 - dD^c(x^{\lambda}) - (1 - x^{\lambda})dD_x^c < 0 \tag{21}$$

Hence here too city size shrinks as the proportion of clannish types in the population increases. Intuitively, as x^{λ} rises, the clan becomes more attractive because all the clannish population is segregated in the clan and they all contribute to the public good. This draws more generalists in the clan, which makes the city even less attractive compared to the clan (because public good provision in the city shrinks). Equilibrium is restored when the idiosyncratic value of the preference parameter δ has risen enough for pivotal individual (i.e. when δ^{γ} is high enough) - by assumption (C1) and by (17) we know that this will eventually happen.

Summarizing the above discussion:

Suppose that (A1)-(A4) hold. Then an equilibrium with segregation in the clan exists if and only if $x^{\lambda} \in (\bar{x}, x^{\mathrm{max}})$, where $1 - \hat{e} > x^{\mathrm{max}} > \bar{x} > 1/2$. In this equilibrium the fraction of each dynasty living in the city is $x^y = G^c(x^{\lambda})$ given by (20), and it varies inversely with x^{λ} over the range $x^y \in (\hat{e}, (1 - \bar{x}))$.

Equilibrium with segregation in the city Finally, consider the equilibrium where the generalist types are segregated in the clan, while the clannish are present in both communities. Repeating the previous steps, the value of δ that leaves type λ indifferent between the clan and the city is:

$$
\begin{aligned}
\delta^\lambda &= q + \lambda - \tau + H(g^n) - H(g^y) \\
&= q + \lambda - \tau + H(\tau d \delta^\lambda x^\lambda) - H[M\tau(1 - x^\lambda) - e]
\end{aligned}
\tag{22}
$$

where the second equation follows from (1), (2), (6), having used

$$
x^{\lambda n} = d\delta^\lambda x^\lambda
\tag{23}
$$

by (6). Again, equation (22) defines a known function $\delta^\lambda = D^e(x^\lambda)$. By the implicit function theorem applied to (22):

$$
\frac{\partial \delta^\lambda}{\partial x^\lambda} \equiv D^e_x = \frac{\tau[d\delta^\lambda H_g(g^n) + M H_g(g^y)]}{1 - dx^\lambda \tau H_g(g^y)} > 0
\tag{24}
$$

where the last inequality follows from (C1). The equilibrium conditions discussed above require that the λ types are present in both the clan and the city. This requires

$$
1/d > D^e(x^\lambda) > 0
\tag{25}
$$

Since $a > /1/d$, (25) then also implies that the generalist types are in all in the city ($\delta^\gamma < 0$).

Note that, by (22) and (13), at the point $x^\lambda = \underline{x}$ we have $D^e(x) = D^s(x) + a = 1/d$ (i.e. all clannish types are in the clan) By (24), then, it follows that condition (25) is satisfied for at least some $x^\lambda \leq \underline{x}$. Since under (A4) we have that $\underline{x} > 0$, we know that an equilibrium with segregation in the clan exists for at least some $x^\lambda < \underline{x}$. Let $x^{\min}$ denote the minimum value of x^λ below which all clannish types are attracted to the city. By (25), $x^{\min}$

is defined implicitly by:

$$0 = q + \lambda - \tau + H(\tau d\delta^\lambda x^{\min}) - H[M\tau(1 - x^{\min}) - e]$$

By By (13) we know that $x^{\min} < \underline{x}$, but we cannot tell whether $x^{\min} > 0$ or $x^{\min} = 0$.

Finally, note that in this equilibrium, the fraction of each dynasty living in the city is

$$x^y = (x^\gamma + x^{\lambda y}) = 1 - dD^e(x^\lambda)x^\lambda \equiv G^e(x^\lambda) \tag{26}$$

Differentiating with respect to x^λ we immediately have that, by (24),

$$G_x^e(x^\lambda) = -dD^e(x^\lambda) - x^\lambda dD_x^e < 0 \tag{27}$$

Hence here too city size shrinks as the proportion of clannish types in the population increases. Intuitively, as x^λ rises, the clan becomes more attractive because more of the clannish are present in the clan and they contribute to the public good. This draws more clannish types from the city into the clan. Equilibrium is restored when the idiosyncratic value of the preference parameter δ has risen enough for pivotal individual (i.e. when δ^λ is high enough) - by assumption (C1) and by (24) we know that this will eventually happen.

Summarizing the above discussion:

Suppose that (A1)-(A4) hold. Then an equilibrium with segregation in the city exists if and only if $x^\lambda \in (x^{\min}, x)$, where $\underline{x} > x^{\min} \geq 0$. In this equilibrium the fraction of each dynasty living in the city is $x^y = G^e(x^\lambda)$ given by (26), and it varies inversely with x^λ over the range $x^y \in ((1-\underline{x}), 1)$.

7.0.5 Proof of Proposition 4

Consider the equilibrium with segregation in the clan. Here equilibrium clan size is $x_t^n = 1 - x_t^{\gamma y} = 1 - G^c(x_t^\lambda)$. Hence (8) can be re-written as:

$$
\begin{aligned}
x_{t+1}^\lambda &= (1 - \bar{p}) + (\bar{p} + \underline{p} - 1)\, x_t^\lambda + (\bar{p} - \underline{p})(1 - G^c(x_t^\lambda)) && (28) \\
&= (1 - \bar{p}) + [2\bar{p} - 1 - (\bar{p} - \underline{p})D^c(x_t^\lambda)]\, x_t^\lambda + (\bar{p} - \underline{p})D^c(x_t^\lambda)
\end{aligned}
$$

where the last equality follows from (20). Denoting by $\delta_s^c \equiv D^c(x_s^\lambda)$ the steady state value of δ^γ in this equilibrium, we can express the steady state fraction of the clannish types by:

$$
x_s^\lambda = \frac{1 - \bar{p} + (\bar{p} - \underline{p})\delta_s^c}{2(1 - \bar{p}) + (\bar{p} - \underline{p})\delta_s^c} \equiv x^{\lambda c} > 1/2 \qquad (29)
$$

where the last inequality follows by noting that in this equilibrium $\delta_s^c > 0$. Thus, in this steady state more than half the population ends up being clannish. The reason is that even some generalists are attracted to the clan, which in turn influences the preferences of their offspring towards the clannish type. Since the steady state is jointly determined by (29) and (15), that implicitly defines $\delta_s^c = D^c(x^\lambda)$, in this equilibrium the steady state is affected by changes in parameters of the static model, since the distribution of types across communities is endogenous, and in turn it influences the evolution of preferences.

Is this steady state stable? Differentiating (28) with respect to x_t^λ, we obtain:

$$
\begin{aligned}
\frac{\partial x_{t+1}^\lambda}{\partial x_t^\lambda} &= (\bar{p} + \underline{p} - 1) - (\bar{p} - \underline{p})\, G_x^c = && (30) \\
&= (\bar{p} + \underline{p} - 1) + (\bar{p} - \underline{p})[(1 - dD^c(x_t^\lambda)) + (1 - x_t^\lambda)dD_x^c]
\end{aligned}
$$

where the second equality follows from (21). Since both terms on the right hand side of (30) are positive, the dynamics is monotonic. If the right-most

term is not too large in the neighborhood of the steady state (or if $\bar{p}-\underline{p}$ is sufficiently small), then the right hand side of (30) is also smaller than unity, so that the steady state is locally stable. Thus, if $x^{\lambda c} \in (\bar{x}, x^{\max})$ defined in the previous subsection, then for any initial condition in this same interval $(\bar{x}, x^{\max})$, the economy remains in the equilibrium with segregation in the clan and eventually reaches the steady state.[50] QED

7.0.6 Proof of Proposition 5

Finally, consider the equilibrium with segregation in the city. Here equilibrium clan size is $x_t^n = 1 - x_t^y = 1 - G^e(x_t^\lambda)$. Hence (8) can be rewritten as:

$$
\begin{aligned}
x_{t+1}^\lambda &= (1 - \bar{p}) + (\bar{p} + \underline{p} - 1)\, x_t^\lambda + (\bar{p} - \underline{p})(1 - G^e(x_t^\lambda)) \qquad (31) \\
&= (1 - \bar{p}) + (\bar{p} + \underline{p} - 1 + (\bar{p} - \underline{p}) d D^e(x_t^\lambda))\, x_t^\lambda
\end{aligned}
$$

where the last equality follows from (26). We can thus express the steady state fraction of clannish types in this equilibrium as

$$
x_s^\lambda = \frac{1 - \bar{p}}{2 - \bar{p} - \underline{p} - (\bar{p} - \underline{p}) d \delta_s^e} \equiv x^{\lambda e} < 1/2 \qquad (32)
$$

where $\delta_s^e \equiv D^e(x_s^\lambda)$ and where the last inequality follows by noting that in this equilibrium $\delta_s^e < 1/d$. Thus, in this steady state less than half the population ends up being clannish. The reason is that some clannish types are attracted to the city, which in turn influences the preferences of their offspring towards the generalist type. Since the steady state is jointly determined by (32) and (22), that implicitly defines $\delta_s^e = D^e(x_s^\lambda)$, in this equilibrium too the

---

[50] Recalling that $(1 - \frac{\hat{e}}{1 - x^{\max}})/d > D^c(x^\lambda) > 0$, and using (29), a sufficient condition for $x^{\lambda c} < x^{\max}$ is

$$
x^{\max} > \frac{1 - \underline{p} - (\bar{p} - \underline{p})\hat{e}}{2 - \bar{p} - \underline{p}}
$$

steady state is affected by changes in parameters of the static model, since the distribution of types across communities is endogenous, and in turn it influences the evolution of preferences.

To assess stability, again differentiate (31) with respect to x_t^λ, to obtain:

$$
\begin{aligned}
\frac{\partial x_{t+1}^\lambda}{\partial x_t^\lambda} &= (\bar{p} + \underline{p} - 1) - (\bar{p} - \underline{p}) \, G_x^e = \\
&= (\bar{p} + \underline{p} - 1) - (\bar{p} - \underline{p}) d[D^e(x_t^\lambda)) + x_t^\lambda D_x^e]
\end{aligned}
\tag{33}
$$

where the second equality follows from (27). If the right-most term is not too large in absolute value in the neighborhood of the steady state (or if $\bar{p} - \underline{p}$ is sufficiently small), then the right hand side of (33) is also smaller than unity, so that the steady state is locally stable. Moreover, if $\bar{p} - \underline{p}$ is sufficiently small, then the right hand side of (33) is also positive, so that the dynamics is also monotonic. Thus, if $x^{\lambda e} \in (x^{\min}, \underline{x})$ defined in the previous subsection, then for any initial condition in this same interval $(x^{\min}, \underline{x})$, the economy remains in the equilibrium with segregation in the city and eventually reaches the steady state. QED

7.1 Data on Urbanization within Europe

Definition of regions   Regions coincide with the current NUTS II definition, with the following exceptions.

In Spain, the Principado de Asturias is aggregated with Cantabria; the Communidad Foral de Navarra is aggregated with La Rioja; the region of Madrid is aggregated with Castilla-la Mancha. Finally, the Canary Islands, the Azores and Madeira have been excluded from the dataset since data on extended family are missing.

In Italy, Piemonte is aggregated with Valle d'Aosta. Trentino Alto Adige, Veneto and Friuli Venezia Giulia have been merged. Umbria is aggregated with Marche. Abruzzo, Molise and Basilicata are merged, although the lat-

ter is geographically disjoint. Note that all these regions are classified by the national statistical institute as belonging to the south of Italy and have similar economic and social conditions. Sicilia and Sardegna, the two insular regions of Italy, have been merged.

In Finland, the regions of South Finland and the Aland Islands have been aggregated.

In Germany, the United Kingdom, the Netherlands and France, regions are aggregated at the NUTS I level. Moreover, further aggregation has been introduced for data consistency. In particular, in Germany the small states of Bremen and Hamburg have been aggregated, though not bordering. Rheinland Pfalz and Saarland have been aggregated, too. In the United Kingdom, the regions of London, East of England, and South East England have been aggregated.

As described in the text, the data on family types are available at the NUTS III level, and were aggregated into the larger NUTS II regions by Duranton and Rodriguez-Pose (2008) by weighting observations by land areas. For consistency, in the above mentioned case where we aggregate NUTS II regions into larger areas, observations are also weighted by land area. Regional area includes water and the source is Eurostat data referring to 1999.

Urbanization Spanish regional population refers to 1591 and is reported by Nadal (1984, p. 74). English regional population in 1600 is reported by Broadberry, Campbel and van Leeuwen (2010, working paper). For Sweden, regional population refers to 1571, it is estimated by Andersson Palm (2000) and can be found online at http://www.tacitus.nu/historisk-atlas/befolkning/lan.htm. Estimates of regional population in 1600 in the Netherlands come from Jan Luiter Van Zanden (forthcoming) and have been kindly provided by the author.

Regional population around 1860 is obtained from the following sources and comes from the University of Utrecht database, available online at http://www.populstat.info,

69

in turn relying on the following sources:

Country	Year	Source
France	1861	French Statistical Institute INSEE, census
Western Germany	1867	Almanach de Gotha, estimate
Italy	1861	Almanach de Gotha, census
Netherlands	1859/60	Almanach de Gotha, census
Portugal	1862	Almanach de Gotha, estimate
Spain	1860	Almanach de Gotha, census
U.K.	1861	Almanach de Gotha, census
Austria	1857	Statistisches Jahrbuch fur Osterreich, census
Ireland	1861	Almanach de Gotha, census
Sweden	1860/1861	Almanach de Gotha, census
Finland	1861	Almanach de Gotha, n. s.

Table 7: Sources. Europe.

7.2   Data on Urbanization in China

Geography: Regions and Counties   There are discrepancies between the two surveys and not all counties in the YB05 are in the GSS06. Six counties had to be matched based on, in particular, the order in which counties are listed and data from previous surveys. Thus, counties 370502, 430421, 441203, 452601, 469007, and 610902 in the GSS were matched with counties 370523, 430412, 441223, 452631, 452723 and 610922 in the YB05 (respectively). In most cases, the matching required replacing a zero with a positive number in the one of the two last digits. This implies, in particular, that the county remained in the same province. In only one case the matching altered the province (469007 vs 452723). The resulting observation is an outlier in favor of our conjecture and it is excluded from our analysis. Our results are robust to using the restricted sample of 69 counties. Villages were extracted from variable qs2 (GSS06) which is a 12 digit code that specifies the province (2 first digit), the county (4), the village (3), and the sub-village (3). The

following table lists the regions, provinces and counties in our sample.

<< Insert table 8 here >>

Variables Rural residency is known from a direct question in the GSS06 (qs3, option 8 'rural') while the YB provides county-level data on total and rural population. Urban Investment: Reported in the YB05 and specifics investments on construction projects involving an investment of at least 500,000 yuan by either individuals, firms, or the government. This includes all investments that take place in county towns and urban areas, investment in construction projects under the direct leadership and management of government agencies at and above county levels and investments by enterprises and institutions at and above county levels are covered in urban investment in fixed assets. Other variables: the GSS reports individual-level education attainment (in years, qb03b) and yearly income (qb12b) from which we constructed village and county means. The YB reports population and land area per-county and we used previous Yearbooks to fill in missing data.

Alternative clannishness measures and regional analysis The results are robust to the three other measures of clannishness we could think of; county average of each village's highest response to the clan organization question (1, 2, or 3), the county average of each village's average response, and the county-level probability that one knows of (and thus arguably a member of) a temple-based clan organization. The results are weakest for the last measure and thus we discuss it here. The sample mean of the probability that one is a member of a temple-based organization is 2.4% implying one organization per 42 people and it is positive in 30% of the counties, and in all regions but the north. The results reported above are robust to using this proxy for clannishness. In repeating model 5 in table 5, for example, the coefficient is -0.25 (significant in 1%) and the R-sq is .32.

Figure 9 highlights that temple-based clan organizations are particularly common in the South Central region. It thus possible that this region distorts the analysis of China as a whole. This however, is not the case. The two upper graphs in figure 10 repeat model 5 (table 5) for the south central region (22 counties) and the rest of China. (The coast variable is added to control for geography in the one region regression.) The two lower graphs repeat the analysis using the probability of being a member of a temple-based clan organization. The coefficient is always negative and significant at the 0.001 or the 0.005 level. The lower right graph reveals, however, an outlier. The coefficient is insignificant after excluding this county (441223 in the Guandong province) but it is remains significant (at the 7.6% level) using the village probability variable

8   References

Andersson Palm L. (2000), "Folkmängden i Sveriges socknar och kommuner 1571-1991" NON TROVO LA CASA EDITRICE!

Bairoch, P., J. Batou and P. Chëvre (1988) La population des villes Européennes de 800 á 1850. Libraire Droz, Genéve.

Bai, C.E., Lu, J., Tao, Z., 2006. "Property Rights Protection and Access to Bank Loans: Evidence from Private Enterprises in China." *Economics of Transition* 14 (4): 611–628.

Ballard, Adolphus. 1913. British Borough Charters 1042-1216. 2010 edition by Cambridge University Press.

Ballard, Adolphus, and James Tait (eds.).1923. British Borough Charters, 1042-1216. Cambridge: Cambridge University Press.

Baten, Joerg, Debin Ma, Stephen Morgan c,d, and Qing Wanga. 2010. "Evolution of Living Standards and Human Capital in China in the

18–20th Centuries: Evidences from Real Wages, Age-heaping, and Anthropometrics." *Explorations in Economic History*, vol. 47: 347–359.

Benabou, Roland and Jean Tirole. 2006. "Belief in a Just World and Redistributive Politics," Quarterly Journal of Economics August Vol. 121, No. 2, pp. 699-746.

Beresford, M. and H. P. R. Finberg. 1973. English Medieval Boroughs: A Handlist. Newton Abbott: David and Charles.

Bian, Yanjie. 1994. "*Guanxi* and the Allocation of Urban Jobs in China." *The China Quarterly*, Vol. 140(Dec.): 971-999.

Bisin, Alberto & Verdier, Thierry, 2000. "A model of cultural transmission, voting and political ideology," European Journal of Political Economy, Elsevier, vol. 16(1), pages 5-29

Bisin, Alberto & Verdier, Thierry, 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences," Journal of Economic Theory, Elsevier, vol. 97(2), pages 298-319

Boerner, Lars and Daniel Quint. 2010. "Medieval Matching Markets." Memo. Free University Berlin

Bowles, Samuel, and Herbert Gintis. 2011. A Cooperative Species. Human Reciprocity and its Evolution. Princeton: Princeton University Press.

Brenner, Robert and Christopher Isett. 2002. "England's Divergence from China's Yangzi Delta: Property Relations, Microeconomics, and Patterns of Development." *The Journal of Asian Studies* 61(2): 609–662.

Broadberry S., Campbell B. and van Leeuwen B. (2010), "English medieval Population: reconciling time series and cross sectional evidence", working paper

73

Broadberry, S. and B. Gupta. 2006. "The Early Modern Great Divergence: Wages, Prices and Economic Development in Europe and Asia, 1500-1800." Economic History Review Vol. 59: 2–31.

Cantoni, Davide and Noam Yuchtman. 2009. "Medieval Universities, Legal Institutions, and the Commercial Revolution". Working paper. Harvard.

Chao, Kang. 1986. Man and Land in China, Stanford: Stanford UP.

Chung, Stephanie Po-yin. 2010. "Chinese Tong as British Trust: Institutional Collisions and Legal Disputes in Urban Hong Kong, 1860s–1980s." *Modern Asian Studies*, 44(6), pp. 1409–1432

Clark, David S. 1987. "The Medieval Origins of Modern Legal Education: Between Church and State." *The American Journal of Comparative Law*, Vol. 35 (4), pp. 653-719.

Clark, Peter. 2009. European Cities and Towns, 400-2000. Oxford: Oxford University Press

Cohen, Myron. 1990. "Lineage Organization in North China," *The Journal of Asian Studies*, 49(3 Aug): 509-534.

Deng, Kent. G. 2000. "A Critical Survey of Recent Research in Chinese Economic History." *Economic History Review*, LIII: 1-28.

Deng, Kent. G. 2010. "Song-Qing China: A Fiscal State or an Anti-Fiscal State?" LSE. Paper presented at a conference on "Fiscal Regimes and the Political Economy of Early States" at Stanford University.

Duranton, Gilles and Andrés Rodriguez-Pose 2009. "Family types and the persistence of regional disparities in Europe", *Economic Geography*, vol 85, January pp 23-47

74

Ebrey, Patricia Buckley and James L. Watson. 1986. *Kinship Organization in Late Imperial China 1000-1940.* Berkeley: University of California Press.

Ekelund, Robert B., Robert D. Tollison, Gary M. Anderson, Robert F. Hebert, and Audrey B. Davidson. 1996. *Sacred Trust: The Medieval Church as an Economic Firm.* Oxford UP.

Fairbank, John King. 1994. China, A New History. Cambridge, MA: Harvard University Presss.

Faure, David. 1989. "The Lineage as a Cultural Invention: The Case of the Pearl River Delta." *Modern China,* vol 15(1): 4-36.

Faure, David. 2007. *Emperor and Ancestor: State and Lineage in South China.* Stanford: Stanford University Press

Fei, John C. H. and Ts'ui-Jung Liu. 1982. "The Growth and Decline of Chinese Family Clans." J*ournal of Interdisciplinary History,* 12(3), pp. 375-408.

Feuerwerker, Albert. 1984. "The State and the Economy in Late Imperial China." *Theory and Society,* vol. 13(3): 297-36.

Freedman, Maurice. 1958. Lineage Organization in Southeastern China. London: Athlone.

Freedman, Maurice. 1966. Chinese Lineage and Society: Fukien and Kwangtung. London: Athlone.

Friedmann, John. 2007. "Reflections on Place and Place-making in the Cities of China." *International Journal of Urban and Regional Research,* 31(2), pp. 257–79

Gellhorn, Walter. 1987. "China's Quest for Legal Modernity." *Journal of Chinese Law,* 1987, 1(1), pp. 1-22.

Goldstone, Jack A. 2008. *Why Europe?: The Rise of the West in World History, 1500-1850.* McGraw-Hill.

Goody, J. 1983. *The Development of the Family and Marriage in Europe.* Cambridge: Cambridge University Press.

Greif, Avner. 1994. "Cultural Beliefs and the Organization of Society." *The Journal of Political Economy*, 102(5), pp. 912-50.

Greif, Avner. 2005. "Commitment, Coercion, and Markets: The Nature and Dynamics of Institutions Supporting Exchange." *Handbook for NIE.* C. Menard and MM. Shirley (eds).

Greif, Avner. 2006. *Institutions and the Path to the Modern Economy.* Cambridge UP

Greif, Avner. 2006a. "Family Structure, Institutions, and Growth." *American Economic Review.* 96(2), pp. 308-312.

Greif, Avner. 2006b. "The Birth of Impersonal Exchange: The Community Responsibility System and Impartial Justice." *Journal of Economic Perspective.* Vol. 20 (2): 221-236.

Greif, Avner, Murat Iyigun, and Diego Sasson. 2011. "Risk, Institutions and Growth: Why England and Not China?" IZA Discussion Paper No. 5598. Available at SSRN: http://ssrn.com/abstract=1801664

Guichard, Pierre and Cuvillier, Jean-Pierre. (1996). Barbarian Europe in Ed. André Burguiére et. al. A History of the Family, Vol. I (Cambridge: Polity Press). 318-378

Guiso, L., P. Sapienza and Luigi Zingales. 2006. "Does Culture Affect Economic Outcomes?" Journal of Economic Perspectives 20: 23-48.

Hamilton, Gary G. 1984. "Patriarchalism, in Imperial China and Western Europe: A Revision of Weber's Sociology of Domination." Theory and Society, Vol. 13(3): 393– 425.

Hamilton, Gary G. 1990. "Patriarchy, Patrimonialism, and Filial Piety: a Comparison of China and Western Europe." *British Journal of Sociology*, 41 (1): 77-104.

Haung, Philip C. 1985. *The Peasant Economy and Social Change in North China*. Stanford: Stanford University Press.

Huang, Youqin. 2006. "Collectivism, Political Control, and Gating in Chinese Cities." *Urban Geography*, 27(6), pp. 507–525.

Herlihy, David. 1985. Medieval Households. Cambridge: Harvard University Press.

Herrmann-Pillath, Carsten. 2009. "Social Capital, Chinese Style: Individualism, Relational Collectivism and the Cultural Embeddedness of the Institutions–performance Link." *China Economic Journal*, Vol. 2(3): 325–350.

Herrmann-Pillath, Carsten. 2010. "What Have We Learnt from 20 Years of Economic Research into Culture?" *International Journal of Cultural Studies*, Vol. 13(4): 317–335.

Horowitz, Richard S. 2002. "Beyond the Marble Boat: The Transformation of the Chinese Military, 1850-1911." Chapter 9 (pp.153-174) in *A Military History of China*. Edited by David A. Grafy and Robin Higham. Cambridge, MA: Westview Press.

Inglehart, Ronald, Miguel Basanez, and Alejandro Moreno. 1998. *Human Values and Beliefs. A Cross-Cultural Sourcebook*. Ann Harbor: Michigan University Press.

Kiser, Edgar and Yong Kai. 2003. "War and Bureaucratization in Qin China: Exploring an Anomalous Case," *American Sociological Review.* 68(4): 511-39.

Khun, Dieter. 2009. *The Age of Confucian Rule. The Song Transformation of China.* Cambridge: Harvard University Press.

Kung, James Kai-sing and Chicheng Ma. 2010. Choosing Not to Vote? The Effect of Clans and Patriarchy on Women's Political Participation in Rural China. Hong Kong University of Science and Technology.

Krausman Ben-Amos, Ilana. 2000. "Gifts and Favors: Informal Support in Early Modern England." *The Journal of Modern History,* Vol. 72, No. 2 (Jun), pp. 295-338.

Kuran, Timur. 1987. Preference Falsification, Policy Continuity and Collective Conservatism. The Economic Journal, 97:642 -65.

Landa, J. T. (1981). A theory of the, ethnically homogeneous middleman group: an institutional alternative to contract law. The Journal of Legal Studies, 10, 348–362.

Landes, DS. 2006. "Why Europe and the West? Why Not China?" *Journal of Economic Perspectives* 20 (2): 3-22.

Langlois, Jr., John D.. 2007. [1998]. "Ming Law." Chapter 3 (pp. 172-220) in *The Cambridge History of China*, vol. 8, p. 2. Edited by Denis Twitchett and John K. Fairbank. Cambridge: Cambridge University Press.

Laslett, Peter. 1969. " Size and Structure of the Household in England Over Three Centuries." Population Studies, Vol. 23, No. 2 (Jul), pp. 199-223

Lau, Chong Chor, Yushend Peng and Shen Chonglin. 1999. "Kinship Networks and Rural Industrialization in Chinese Villages." In *China Review*, edited by Chong Chor Lau and Xiao Geng. Chinese University Press, pp. 197-212.

Lau, Lawrence J., Yingi Qian, and Gérard Roland. 2000. "Reform Without Losers: An Interpretation of China's Dual-Track Approach to Transition." *Journal of Political Economy* 108 (1): 120-43.

Levi, Margaret. 1997. Consent, Dissent, and Patriotism. Cambridge: Cambridge University Press.

Li, Grace. 2009. "The PRC Contract Law and Its Unique Notion of Subrogation." *Journal of International Commercial Law and Technology*, Vol. 4(1): 12-21.

Li, Hongbin, Lingsheng Meng c, Qian Wang d, Li-An Zhou. 2008. "Political Connections, Financing and Firm Performance: Evidence from Chinese Private Firms." *Journal of Development Economics*, Vol 87(2): 283–99.

Liangqun, Liu and Murphy, Rachel. 2006. "Lineage Networks, Land Conflicts and Rural Migration in Late Socialist China." *Journal of Peasant Studies*, Vol 33(4): 612-45.

Lin, Justin Yifu. 1995. "The Needham Puzzle: Why the Industrial Revolution Did Not Originate in China." *Economic Development and Cultural Change*, Vol. 43(2): 269-92.

Liu, Hui-Chen Wang. 1959. "The Traditional Chinese Clan Rules." *Monographs of the Association for Asian Studies*, vol. 7. New York: J. J. Austion

Liu, Kwang-Ching. 1988. "Chinese Merchant Guilds: An Historical Inquiry." *The Pacific Historical Review*, Vol. 57(1): 1-23.

Lu, Jiangyong and Zhigang Tao. 2010. "Determinants of Entrepreneurial Activities in China." *Journal of Business Venturing*, Vol 25(3): 261 -273.

Lu, Jiangyong, Zhigang Tao, and Zhi Yang. 2010. "The Costs and Benefits of Government Control: Evidence from China's Collectively-owned Enterprises." *China Economic Review*, Vol 21(2): 282 -92.

Ma, Debin. 2004. "Growth, Institutions and Knowledge: a Review and Reflection on 18th-20th Century Chinese Historiography." *Australian Economic History Review*, 44(3), pp. 259-77.

Ma, Debin. 2007. "Law and Economic Growth: the Case of Traditional China- a Review with Some Preliminary Hypotheses." Working Paper. LSE.

Ma, Debin. 2008. "Incentives and Information: an Institutional Perspective on the Traditional Chinese State and Great Divergence in the Early Modern era." Working Paper. LSE.

Maddison, Angus. 2007. Development Centre Studies Chinese Economic Performance in the Long Run - Second Edition, Revised and Updated: 960-2030 AD (Development Centre Studies).

Mann, Susan. 2002. "Women, Families, and Gender Relations." Chapter 8 (pp. 428-472) in *The Cambridge History of China*, vol. 8. Edited by Denis Twitchett and John K. Fairbank. Cambridge: Cambridge University Press.

McEvedy C. and Jones R. (1978): *Atlas of World Population History.* Penguin Reference Books (Paperback)

McKnight, Brian E. . 1985. "Song Legal Privileges." J*ournal of the American Oriental Society*, Vol. 105(1): 95-106.

Medieval Commune. 1911. In Encyclopædia Britannica (on the Public Domain). ncyhttp://www.1911encyclopedia.org/Medieval_Commune. Accessed Feb. 5, 2007.

Meyer, Ramon H. and Yeh-chien Wang. 2002. "Economic Developments, 1644–1800." Chapter 10 (pp. 563-646) in *The Cambridge History of China*, vol. 9. Edited by Denis Twitchett and John K. Fairbank. Cambridge: Cambridge University Press.

Mitterauer, Michael, and Reinhard Sieder. 1982. The European Family. Oxford: Basil Blackwell.

Mocarelli, Luca. 2008. "Guilds Reappraised: Italy in the Early Modern Period." *International Review of Social History*, 53, pp. 159–178.

Moll-Murata, Christine. 2008. "Chinese Guilds from the Seventeenth to the Twentieth Centuries: An Overview." *International Review of Social History*, 53, pp. 213–47.

Nadal J. (1984), "La poblaciòn española (siglos XVI a XX)", Ariel

Nader, Helen. 1990. Liberty in Absolutist Spain. The Habsburg Sale of Towns, 1516-1700 (Baltimore and London: The Johns Hopkins University Press.

Nakamura, Shigeo. 2004. "Was Traditional Chinese Law a Mere Model? Part Two," *International Journal of Asian Studies*, 1(2): 297-322.

Nee, Victor. 1992. "Organizational Dynamics of Market Transition: Hybrid Forms, Property Rights, and Mixed Economy in China." *Administrative Science Quarterly*, Vol. 37(1): -27.

Nee, Victor. 2010. "Bottom-up Economic Development and the Role of the State." Cornell University CSES Working Paper #49. Forthcoming, Sociologica.

Osborne, Ann, 2004: "Property, Taxes, and State Protection of Rights," in Madeleine Zelin et al., eds., *Contract and Property in Early Modern China*, Stanford, California: Stanford University Press, pp. 120-158.

Peng, Yusheng. 2004. "Kinship Networks and Entrepreneurs in China's Transitional Economy." *American Journal of Sociology*, Volume 109 (5): 1045–74.

Perdue, Peter, 1987: *Exhausting the Earth: State and Peasant in Hunan, 1500-1850.* Cambridge, Mass.: Harvard University Press.

Pirenne, Henri. 1969. *Medieval Cities: Their Origins and the Revival of Trade.* Princeton: Princeton University Press.

Platteau, Jean-Philippe. 2000. Institutions, Social Norms and Economic Development. Amsterdam, the Netherlands: Harwood Academic Press.

Pomeranz, Kenneth. 2000. *The Great Divergence: China, Europe and the Making of the Modern World Economy.* Princeton, NJ: Princeton University Press

Potter, Pitman B., 1999. "The Chinese Legal System: Continuing Commitment to the Primacy of State Power." *The China Quarterly*, No. 159(Special Issue): 673-683.

Putnam, Robert D. 1993. Making Democracy Work. Princeton, NJ: Princeton University Press.

Pyatt, T.R., Redding, S.G. (2000), "Trust and forbearance in ethnic Chinese business relationships in Hong Kong and Thailand." *Journal of Asian Business*, 16(1), pp. 41-63.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." American Economic Review 83 (5): 1281-1302.

Razi Z.. 1993. The Myth of the Immutable English Family. Past & Present 140(Aug), Pp. 3-44

Redding, Gordon and Michael A. Witt. 2009. "China's Business System and its Future Trajectory." *Asia Pacific Journal of Management*, 26: 381-99.

Redding, S. Gordon. 1993. *The Spirit of Capitalism.* Berlin: Gruyter.

Robinson, David. M. 2000. "Banditry and the Subversion of State Authority in China: the Capital Region during the Middle Ming Period (1450-1525)." *Journal of Social History*, Spring, pp. 527-63.

Roland, Gérard. 2004. "Understanding institutional change: fast-moving and slow-moving institutions," Studies in Comparative International Development 38(4), p. 109-131.

Rosenthal Jean-Laurent and R. Bin Wong. 2011. *Before and Beyond Divergence. The Politics of Economic Change in China and Europe.* Harvard University Press.

Rowe, William T. 1984. Hankow: Commerce and Society in A Chinese City, 1796-1889.

Rowe, William T. 2002. "Social Stability and Social Change." Chapter 9 (pp. 473-562) in *The Cambridge History of China*, vol. 9. Edited by Denis Twitchett and John K. Fairbank. Cambridge: Cambridge University Press.

Ruskola, Teemu. 2000. "Conceptualizing Corporations and Kinship: Comparative Law and Development Theory in a Chinese Perspective," *Stanford Law Review.* 52(6 Jul.): 1599-1729.

Shiue, Carol H. and Wolfgang Keller. 2007. "Markets in China and Europe on the Eve of the Industrial Revolution." *American Economic Review*, 97(4), pp. 1189-1216.

Schofield, Phillipp R. 2003, Peasant and community in medieval England, 1200-1500. MacMillan. New York.

Skinner, William G. 2001 [1974]. *Marketing and Social Structure in Rural China*. Ann Arbor, MI: The Association of Asian Studies.

Smith, Joanna F. Handlin. 1987. "Benevolent Societies: The Reshaping of Charity During the Late Ming and Early Ch'ing." *The Journal of Asian Studies*, Vol. 46, No. 2 (May, 1987), pp. 309-337.

Sommer, Matthew H. 2007. "The Adjudication of Wife-selling in Qing County Courts: 272 Cases from Ba, Nanbu, and Baodi Counties." Memo. Stanford University.

Szonyi, Michael. 2002. *Practicing Kinship. Lineage and Descent in Late Imperial China*. Stanford: Stanford University Press.

Su, Jun and Jia He. 2010. "Does Giving Lead to Getting? Evidence from Chinese Private Enterprises." *Journal of Business Ethics*, Vol 93(1): 73 -90.

Tabellini, Guido. 2008. "The Scope of Cooperation." *Quarterly Journal of Economics*, 123, pp. 905-50.

Tabellini, Guido. 2008a. "Institutions and Culture." Presidential Address at the European Economic Association in Budapest, Journal of the European Economic Association, forthcoming.

Todd, Emanuel 1990 : "*L'invention de l'Europe*", Paris, Seuil

Thøgersen, Stig. 2002. "Book Review: Village Economy and Culture in Contemporary China." *Modern China*, Vol. 28: 253-74.

Tsai, Lily L., 2007. *Accountability without Democracy.* Cambridge: Cambridge University Press.

Yu, Guanghua and Hao Zhang. 2008. "Adaptive Efficiency and Financial Development in China: the Role of Contracts and Contractual Enforcement." *Journal of International Economic Law*, Vol 11(2): 459 -494.

Voigtländer, Nico and Hans-Joachim Voth. 2009. "The Three Horsemen of Growth: Plague, War and Urbanization in Early Modern Europe." CEPR Discussion Paper No. DP7275.

Watson, James L. 1982. "Chinese Kinship Reconsidered: Anthropological Perspectives on Historical Research." *The China Quarterly*, No. 92. (Dec.), pp. 589-622.

Weber, Max. 2007. *General Economic History.* Cosimo Classics

Whyte, Martin K. 1995. "The Social Roots of China's Economic Development." *China Quarterly* 144:999–1019.

Whyte, Martin K. 1996. "The Chinese Family and Economic Development: Development: Obstacle or Engine?" *Economic Development and Cultural Change*, 45 (1): 1–30.

Williams, Michael. 2000. "Dark Ages and Dark Areas: Global Deforestation in the Deep Past," *Journal of Historical Geography*, 26(1), pp. 28–46.

Zelin, Madeleine. 2004. "Managing Multiple Ownership at the Zigong Salt Yard," in Zelin, K.Ocko and Gardella, eds., *Contract and Property in Early Modern China*: Stanford: Stanford University Press.

Zelin, Madeleine. 2009. "The Firm in Early Modern China." *Journal of Economic Behavior & Organization*, Vol 71(3): 623–37.

Zelin, Madeleine, et al., eds., 2004. *Contract and Property in Early Modern China*, Stanford, California: Stanford University Press.

Zhang, Mo. 2008. "From Public to Private: The Newly Enacted Chinese Property Law and the Protection of Property Rights in China." *Berkeley Business Law Journal*, Vol. 5(2): 317-63.

Zhenman, Zheng. 2001. *Family Lineage Organization and Social Change in Ming and Qing Fujian.* Translated by Michael Szonyi. Honolulu: University of Hawaii Press.
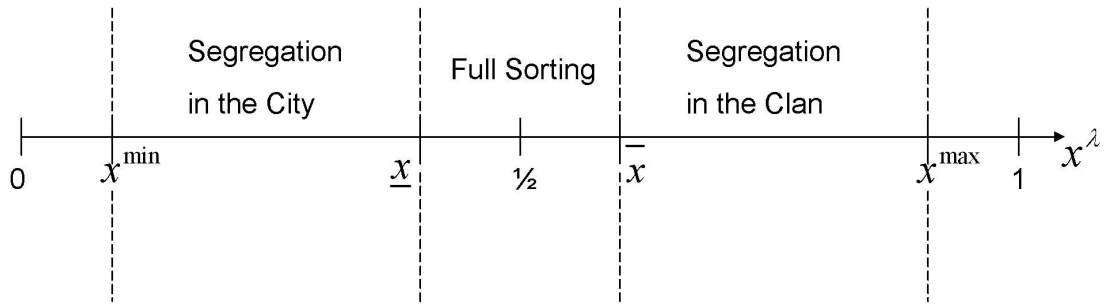
Figure 1: Equilibrium Configurations
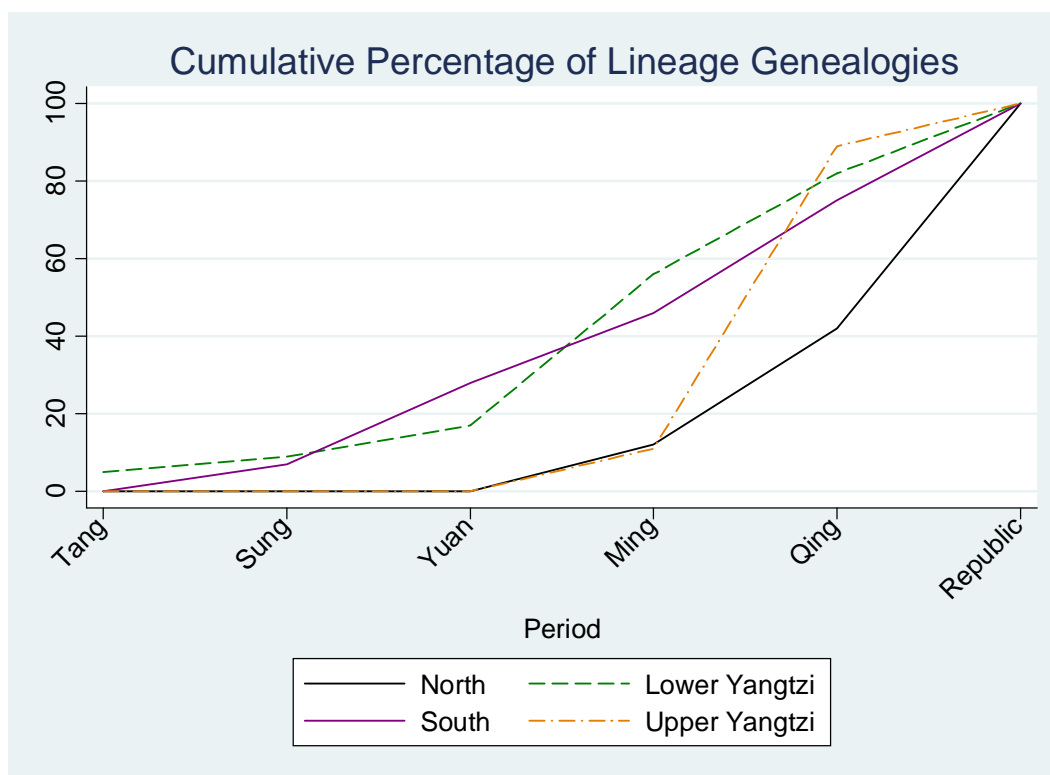
Figure 2: Cumulative Percentage of Lineage Genealogies

Source: We thank Carol H. Shiue for sharing this data with us. (From the "Survey of Social Demographic Data," by Telford).
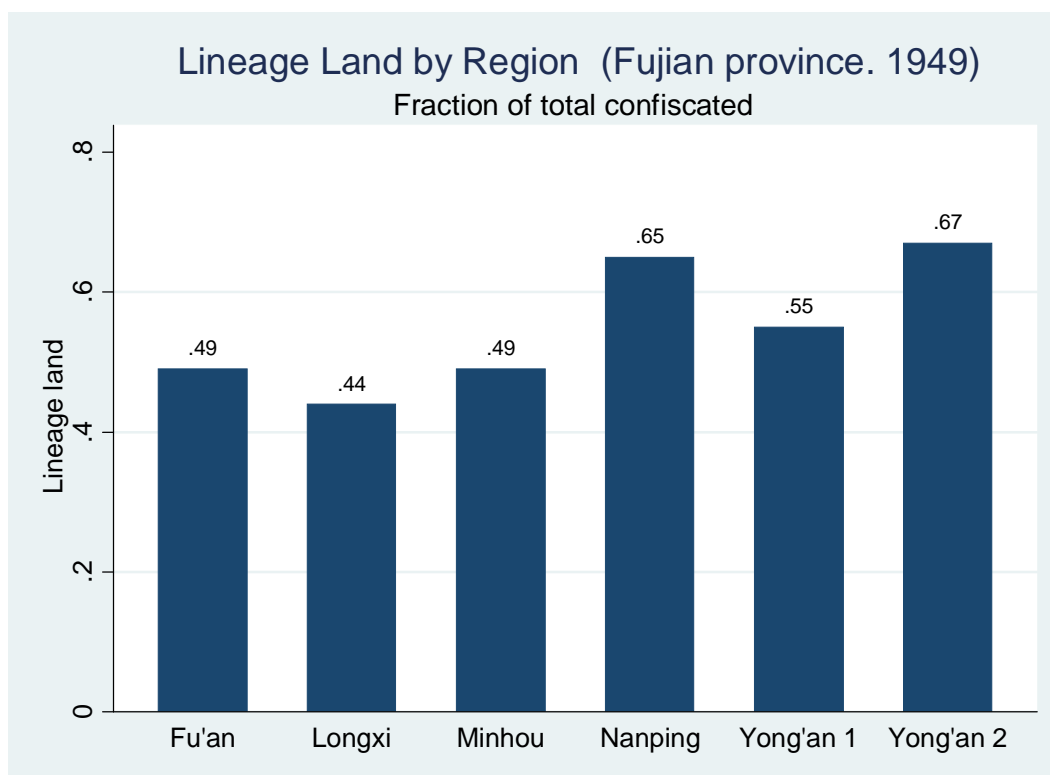
Figure 3: Lineage in the Fujian Province, 1949
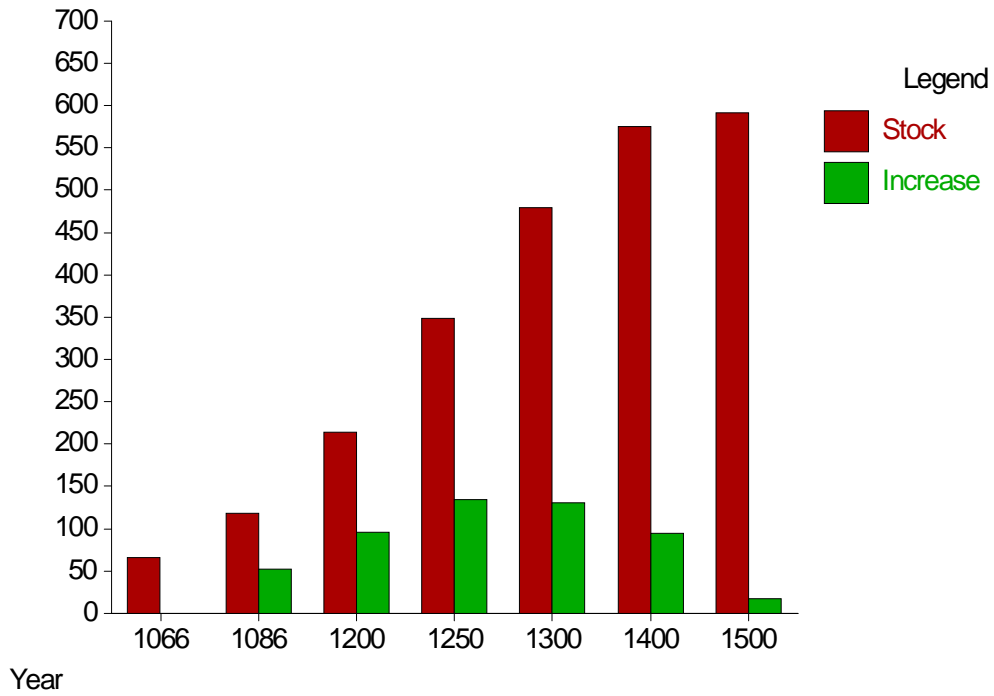
Source: Zhenman, 2001, table 5.1.

English Boroughs

Figure 4a: Urban Growth, England
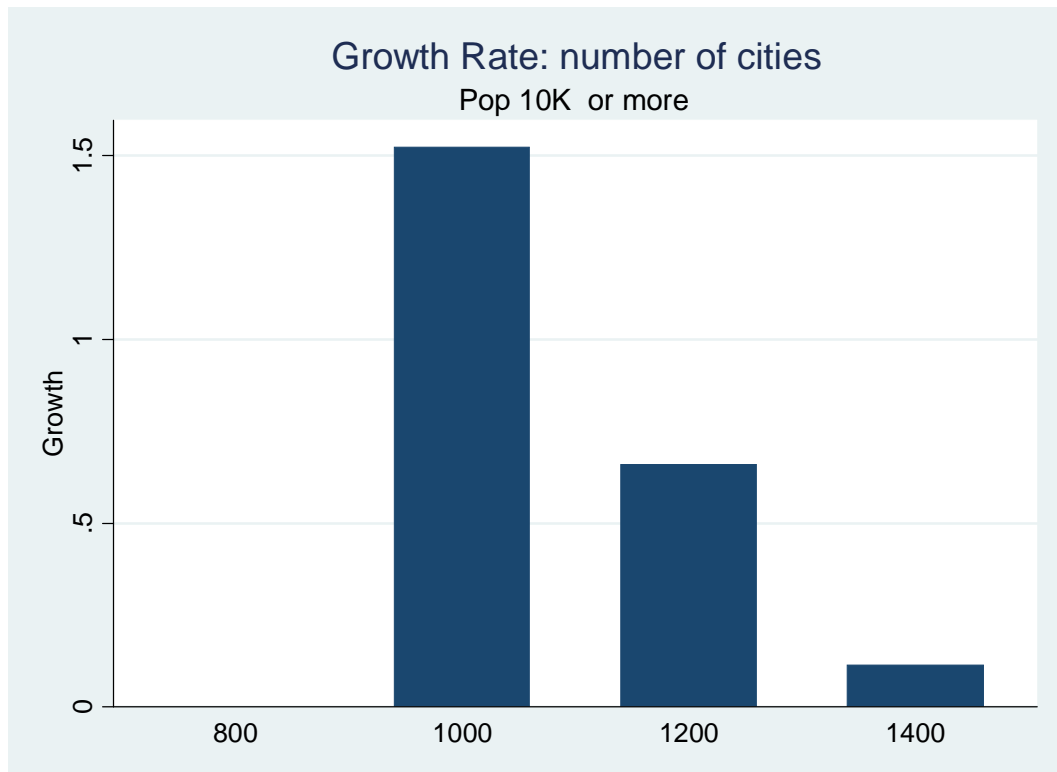Source: Bersford and Finberg (1973).

Figure 4b: Urban Growth, Europe

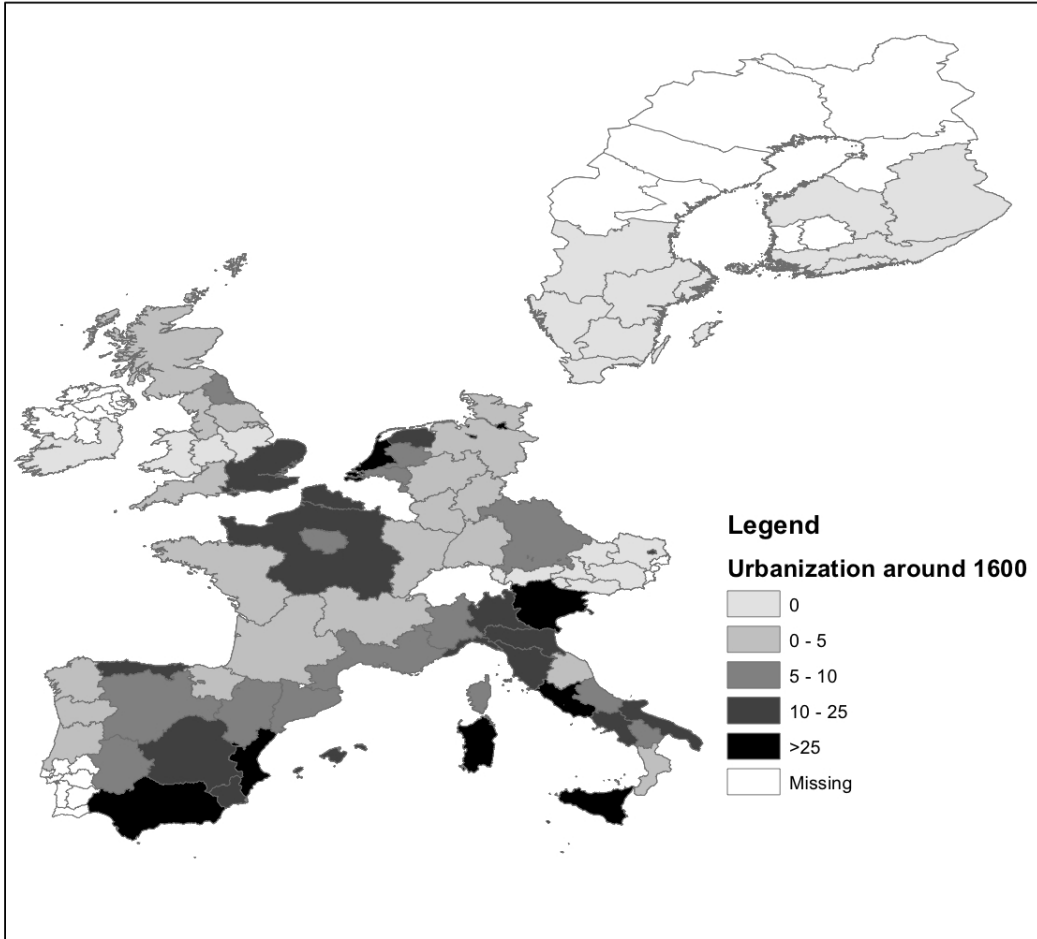Source: Bairoch, Paul, Jean Batou and Pierre Chèvre (1988). Excluding Iberia.

Figure 5: Urbanization around 1600

Note: 7 observations are constituted by disjoint regions in the map. They are: Bremen Hamburg, Asturias-Cantabria, Mediterrean Fr, Abruzzo-Molise-Basilicata, Sicilia-Sardegna, Smaland med oarna, Scotland.
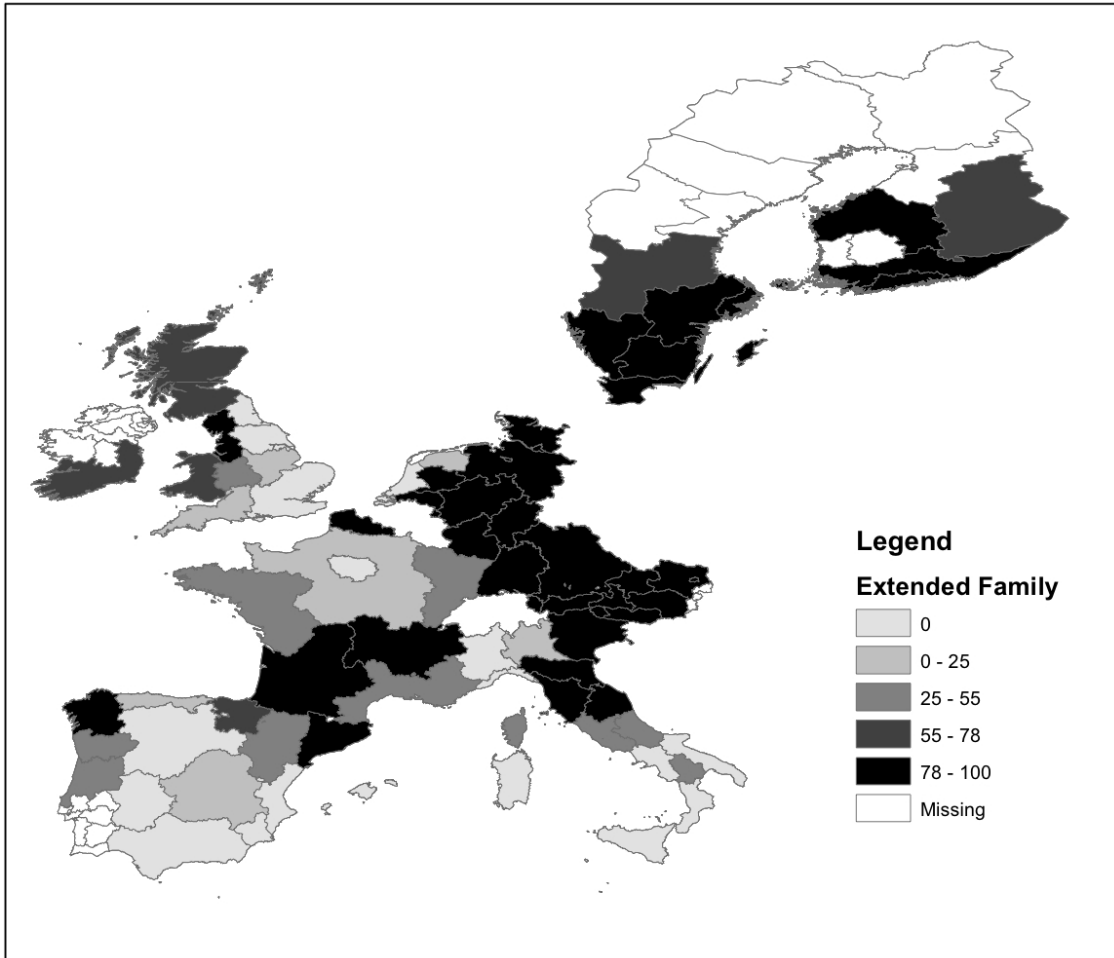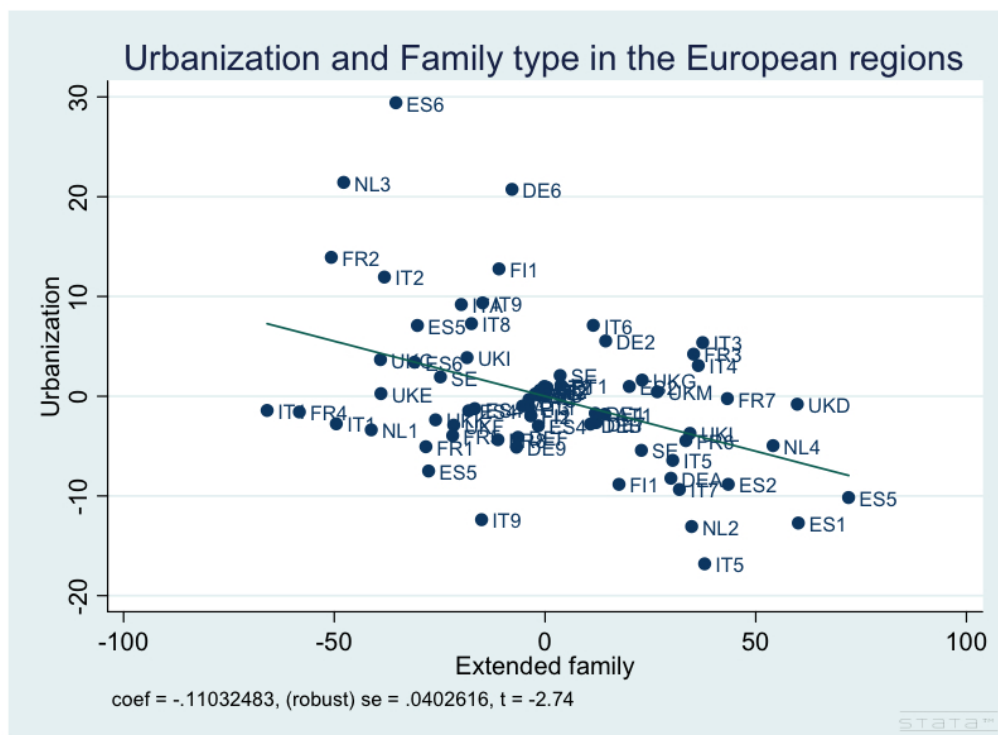
Figure 6: Extended Family

Figure 7: Urbanization and Family type in the European regions

Note: 7 observations are constituted by disjoint regions in the map. They
are: Bremen Hamburg, Asturias-Cantabria, Mediterrean Fr, Abruzzo-Molise-
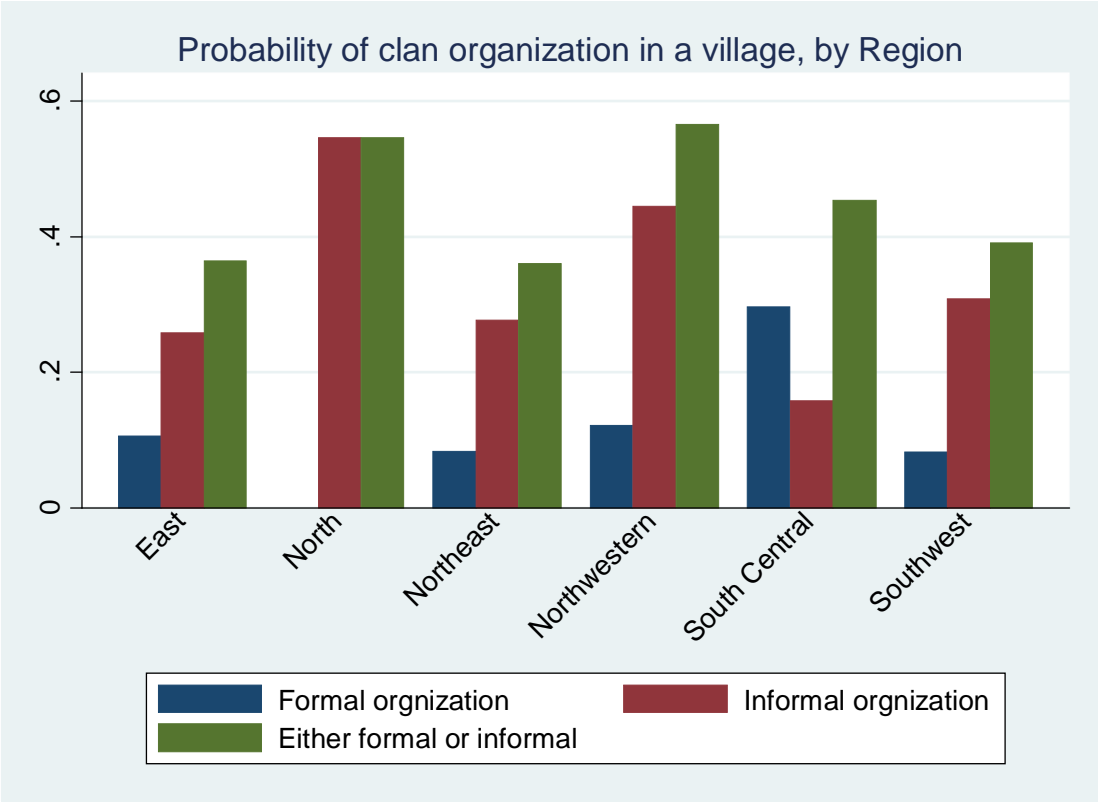Basilicata, Sicilia-Sardegna, Smaland med oarna, Scotland.

94

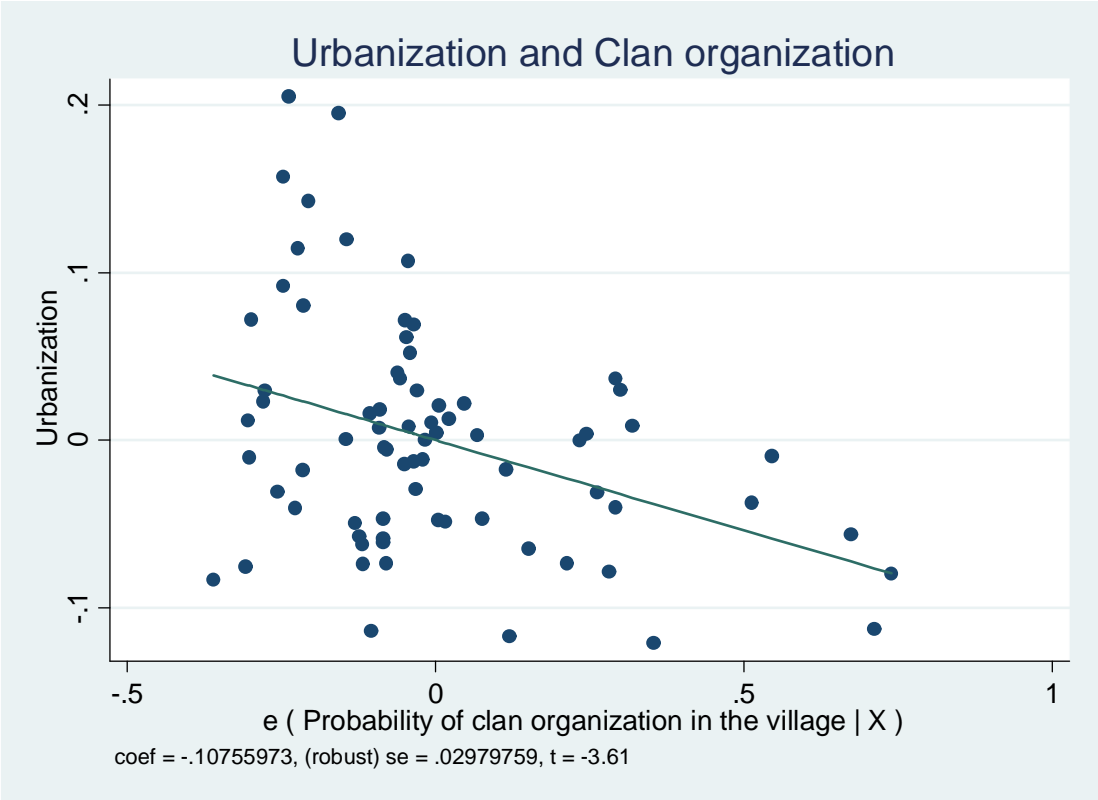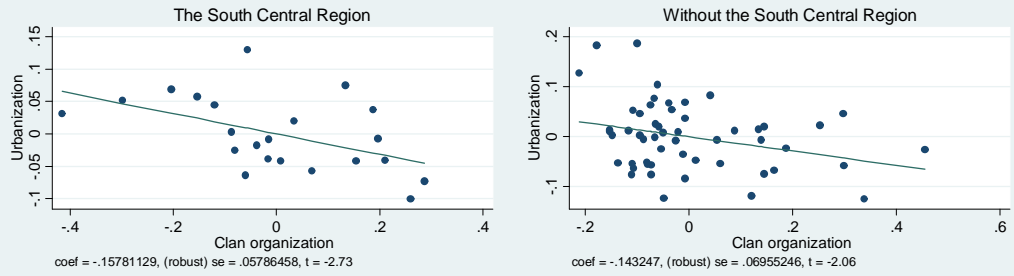Figure 8: Clan Organizations by Regions. China.

Figure 9: Urbanization and Clan Organization. China.
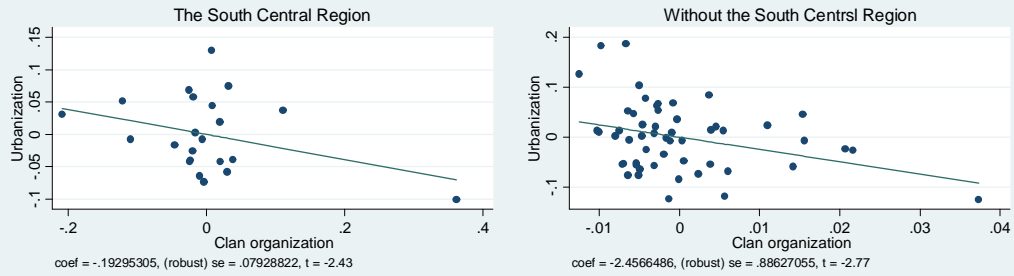
Figure 10: Urbanization and Clan Organization, Rubostness. China.

Dependent variable	(1)	(2)	(3)	(4)
	Urbanization around 1600			
Extended family	-0.118***	-0.124**	-0.120**	-0.110**
	(0.035)	(0.040)	(0.041)	(0.040)
Sea	11.709***	14.476**	14.966**	14.800**
	(2.076)	(5.476)	(5.459)	(5.947)
Atlantic Ocean	-3.995*	-7.130	-7.420	-7.716
	(2.177)	(5.878)	(5.942)	(7.068)
Longitude		-0.407	-0.463	-0.442
		(0.375)	(0.359)	(0.361)
Population density			14.545***	8.074
			(4.095)	(6.894)
Capital				5.546
				(4.221)
Observations	74	74	74	74
Adjusted R-squared	0.467	0.480	0.499	0.516

Robust standard errors in parentheses, clustered by country
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Regression Results, Europe

Variable	Source	N	Mean	Min	Max	Variance
Urbanization	YB05	76	0.156	0	0.444[51]	0.014
Any Clanorg	GSS06	76	0.426	0	1	0.150
Informal Clanorg	GSS06	76	0.277	0	1	0.127
Formal Clanorg	GSS06	76	0.149	0	1	0.077
Education	GSS06	76	7.605	3.450	12.550	2.531
Income	GSS06	76	0.045	0.017	0.162	0.001
Ln Density	YB05	76	3.234	-0.55	4.978	1.092
Urban Fixed Invest	YB05	76	1.473	0.092	12.266	3.440

Table 4: Summary Statistics, China

---

[51]Excluding one outlier with rate of 0.903. See below.

Dependent variable	(1)	(2)	(3)	(4)	(5)
			Urbanization Rate		
Clan organization	-0.084***	-0.106***	-0.106***	-0.105***	-0.108***
	(0.025)	(0.026)	(0.025)	(0.026)	(0.03)
Education	0.006	-0.002	0.000	-0.001	-0.001
	(0.008)	(0.008)	(0.01)	(0.008)	(0.008)
Rural HH income		0.932**	0.949**	0.529	0.614
		(0.407)	(0.388)	(0.418)	(0.408)
Population density			-0.021	-0.028**	-0.021**
			(0.013)	(0.013)	(0.010)
Urban fixed investment				0.012***	0.014***
				(0.004)	(0.004)
Regional fix effects	No	No	No	No	Yes
Obs.	75	75	75	75	75
R-squared	0.095	0.165	0.237	0.283	0.377

Standard in parentheses
***p<0.001. **p<0.05, *p<0.1

Table 5: Regression Results, China

Clan organization and cooperation

Role	N	Percentage
Dispute resolution in village	198	71.48%
Relations with other villages	166	59.93%
Contributes to public welfare	156	56.32%
Provide public goods in clan	153	55.23%
Relations with government	135	48.74%
Assist village committee	135	48.74%
Provide job information	131	47.29%
Assist the government	127	45.85%
Promote culture	124	44.77%
Assist in economic activity	113	40.79%

Table 6: The Roles of Clan Organizations.
Source: qf13 in GSS06.

Region	Province	County	Region	Province	County
North	Hebei	130128	**South Central**	Henan	411224
		130282			411321
		130621			411626
		131024			420322
	Shanxi	140622			420324
		142301			420582
	Nei-Mongol	150429		Guangdong	422826
Northeast	Liaoning	210922			430221
		211221			430421
		211422			430724
	Jilin	220722			431230
	Heilongjiang	230381		Guangxi	440681
		231223			441223
East	Jiangsu	320324			441424
		320381		Hainan	441823
		320623	**Southwest**	Chongqing	445221
		320826		Sichuan	450922
		321121			452631
	Zhejiang	330182			452632
		330329			452723
		330902			500227
		331123			510525
	Anhui	340421		Guizhou	510683
		340828			510923
		341324			511524
		341881		Yunnan	513224
	Fujian	350425			513423
		350725			520330
	Jiangxi	360428		Shaanxi	522323
		361122			522701
	Shandong	370523			530421
		370828		Gansu	532301
		370883			532527
		371424			610323
		371624			610431
South Central	Henan	410481			610922
		410621			622103
		410922			622621

Figure 1:

Table 8: Regions, Provinces, and Counties.