

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; This version: July 2012

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 groups with which individuals identify are basic units of cooperation. Such groups impact institutional development because intra-group moral commitment reduces enforcement cost implying a comparative advantage in pursuing collective actions. Moral groups perpetuate due to positive feedbacks between morality, institutions, and the implied pattern of cooperation

*Stanford University, Stanford, 94305, CA, USA (Avner@stanford.edu) and Università' L. Bocconi. Via Sarfatti, 25. 20136 Milano, Italy (Guido.Tabellini@unibocconi.it). We thank CIFAR and the ERC for financial support, Chen Cheng, Alessandra Fenizia, Claudio Sormani and Edoardo Teso for excellent research assistance, and Marco Croci, Carsten Herrmann-Pillath, Debin Ma, James Kai-sing Kung, Roy Mill, Joel Mokyr, Carol Shiue, Yannay Spitzer, and Yiqing Xing for comments. We particularly benefitted from the generosity of Ying Bai and James Kai-sing Kung who directed us to the data regarding China, of Gilles Duranton and Andrés Rodríguez-Pose who gave us their data on family types, and of Guido Alfani for help with historical data on regional population in Europe.

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 was self-governed cities, whose members were drawn from many kinship groups. 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 Song Dynasty (960-1279). A clan is a common descent group consisting of families tracing their patrilineal descent back to one common ancestor who settled in a given locality. 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.

How was cooperation sustained in these different organizations? Why did 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 and institutional arrangements in modern Western and Chinese societies? These are the questions addressed in this paper.

In any social organization, cooperation is sustained by a combination of intrinsic and extrinsic motivation. Individuals are socialized by their parents and peers to identify with a social group and this creates moral obligations. But cooperation is also influenced by material rewards or punishments provided by institutions. 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.¹

Institutional mechanisms also differ between the clan and the city: clan enforcement mainly relies on informal institutions, whereas the city relies more 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.

After a brief historical review, the first part of this paper investigates a theoretical model that highlights the cultural and institutional differences between clan and city. A first result concerns the impact of morality on organizational forms. The prevalence of one or the other organizational form depends on the distribution of values in society. Large clans are more likely if limited morality is widespread, whereas a city is more likely to flourish if moral obligations are not kin based. A second result concerns the impact of social organization on

¹See Tabellini (2008) for a literature review and a contribution. See Bisin and Verdier (2001), Guiso, et al (2006) and Greif (2012) for recent empirical and theoretical economic analysis of morality.

morality. We study how morality evolves to reflect prevailing social arrangements and show that 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 reveals the mechanism behind 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 provides public goods, the share of the population living in the city is small, and intra-city institutions are weak. This equilibrium captures the arrangements that prevailed in China.

Conversely, if generalized morality is initially widespread, the organization of society moves to an 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. 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 social, 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 modern China, despite the revolution backlash under the Chinese communist regime and its hostility towards clan-based organizations.

Cultural traits and organizational forms vary also within Europe, and not just between Europe and China. In the final part of the paper 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 within Europe reflect these different family traditions, with early urbanization being much more diffused in the European regions that had weak family ties.

Our explanation of their cultural and institutional bifurcations also contributes to the debate of why Europe and China followed 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, technologies, and institutions from the East to West.²

Our analysis provides analytical foundations to the structural view, highlighting that complementarity between culture, social organizations, and institutions can lead to persis-

²See Deng (2000) and Brandt, Ma, and Rawski (2011) for reviews. Recent contributions include Pomeranz (2000), Brenner and Isett (2002), Voigtländer and Voth (2009), and Rosenthal and Wong (2011).

tent cultural and institutional bifurcation. Yet, it also highlights that the city and clan have distinct comparative advantages and recognizes that culture is endogenous and responsive to environmental changes. Thus, while the debate regarding the West and China currently focuses on either economic and political outcomes or political economy explanations, this paper highlights the importance of distinct institutional and cultural foundations. These different foundations can support comparable outcomes yet imply distinct long run development trajectories.

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 (or organizations) are institutionally important because they constitute basic units of cooperation. In China and Europe, for example, distinct social organizations led to differences in the institutional foundations of the market, poor relief, and state administration, among other distinctions. Second, social groups are also held together by mutual moral obligations and group-specific morality, not just by economic interests or other common attributes. Third, although very persistent, the forces that glue individuals to a 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.

2 Historical Background

This section highlights the historical and cultural background that motivates our analysis. The late medieval period was a turning point in the organization of cooperation in China and Europe. The "clan as a Chinese institution .. have prevailed some 800 year, beginning with the Song dynasty [(960-1297)]" while "the birth of cities" in Europe occurred around the same time.³ In China and Europe people self-organized themselves in choosing new settlements during a prolonged period of intense migration. Yet, they organized themselves differently in a process set in motion by distinct initial morality. In each society individuals made different choices regarding whether to resettle in the company of their kinship group.

2.1 How cooperation was supported: the clan and the city

In China 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 hallmark of the emergence of clans during the Song Dynasty (960-1297), is that commoners began keeping genealogies. At that time, European commoners were adopting surnames for the first time. In sharp contrast to China the most common surnames do not designate common-descent (e.g., Smith, Clark, Draper, Taylor).

Because genealogies in China are associated with clans, the distribution of genealogies over time and regions proxies the distribution of clans. The most detailed analysis of genealogies we know of is Telford (1986), that provides information regarding over 1,000 genealogies.

³Fei and Liu (1982), p. 393; Pirenne (1969), p. 213. A clan is composed of lineages that are descent groups with shared assets. We use the terms interchangeably.

The limitations of Telford’s data are many. The sampling process is not reported and, as Telford noted, which genealogies survived and why is unclear. Moreover, we don’t know whether a clan’s first ancestor is fictional or real.

Table 1 reveals that clans first emerged in the east and south — areas that attracted migration during the Song (960-1279) — and not in the north, the birthplace of the Han people that during this period were out-migrating to the East and South. The table presents the percentage of each region’s genealogies – records of a clan’s members – that trace the clan’s origin to a given period. Thus, for example, about 40 percent of the sample genealogies from the east trace the clan’s origin to a year prior to 1644. The data thus suggests that prior to 1368, clans prevailed in the east and south. They emerged in the north and west only by the Ming period. This period witnessed a large migration back to the north and west, following depopulation due to the Yuan-Ming and the Ming-Qing westward expansion. These inter-regional differences in genealogies do not seem to reflect differences in population density as the regional shares of clans and population grew closer over time (see Heijdra, 1998, p. 440).

<< Insert table 1 around here >>

The clans, headed by their elders, performed “many functions related to education, ceremony, social security, and maintenance of law and order. Whereas in a modern society, most of these functions are performed by the government, the clan was a primary social group (or organization) through which these functions were carried out” in pre-modern China (Fei and Liu, 1982, p. 375). Clans created charitable trusts that financed religious, educational, and relief services to their members. Ancestral halls served as a meeting place for clans’ leaders who, among other roles, resolved disputes among members (Ruskola, 2000, p. 1631).

In Europe instead, settlements were created by and attracted individuals unrelated by blood. Urban revival transpired from the 10th to the 14th centuries as the number and size of cities substantially increased. For example, the number of chartered cities rose in England by a factor of 12 from 1066 to 1600.⁴

From the late medieval period, cities were the loci of cooperation among their members. Most European cities had independent administrative, financial, judicial, and military capacities and provided local public goods such as protection, contract enforcement, and social safety nets (e.g., Pirenne, 1969). Rural settlements were similarly organized in being neither composed nor governed along kinship lines.

There were, obviously, cities in China but the urbanization rate was lower and relatively constant. China’s urbanization rate (cities bigger than 10,000) remained between three to four percent from the 11th to the 19th century while Europe advanced from zero to 10 percent (Table 2). As late as 1900, when more than 33 percent of the European population lived in cities with a population of at least 5,000, only 14 percent of the Chinese population lived in such cities (Wood, 2003, tables 1, 2).

<<insert table 2 around here >>

⁴Bairoch, et al (1988), Bersford and Finberg (1973).

Moreover, the cities in China were not units of cooperation. Prior to the 19th century, large Chinese cities were centers of administration and of exchange whose members did not mutually and jointly provide significant local public goods. Small Chinese cities (townships) were often dominated by the local clans (Skinner, 2001). In Europe, cities were providers of local public goods and, to the best of our knowledge, small cities were, in general, not dominated by kin groups.

2.2 The origins of different value systems

Medieval China and Europe also had very different value systems and religious beliefs.⁵ In China, the cultural background is that of Neo-Confucianism reviving the importance of moral obligations among kin as the basis for social order and placing obligations to kin above obligations to the law. Consider the following discussion in Confucius' *Analects*. "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).

The morality of obligations to kin, and the associated solidarity among kin, came under attack early in the first millennium by the advance of Buddhism that emphasized moral obligations towards the religious community. For political reasons, however, the Tang Dynasty turned to persecuting Buddhism during the 9th century (Kuhn, 2009, pp. 14-5) and in response Buddhism also began advocating moral obligations toward kin in China.⁶ Moreover, in response to Buddhism, Neo-Confucianism was formulated around that time, and it synthesizes Taoist cosmology and Buddhist spirituality around the core of Confucian concern with kinship.

In contrast, Christianity has fostered the dissolution of kinship groups and promoted generalized morality. Initially, following the fall of the Roman Empire, kin relations and limited morality were strengthened by large-scale immigration of Germanic and other tribes. The importance of kinship is reflected, for example, in the Germanic legal codes according to which membership in a large kinship group was necessary for legal rights. Yet, tribalism declined in Europe particularly due to the Church's marriage dogma that restricted practices such as polygamy and kin marriage (Goody, 1983). The Church also undermined moral obligations among kin by emphasizing salvation of the individual, the community of believers, and generalized morality. By the 10th century, large kinship groups remained important only on Europe's social and geographical margins (e.g., high nobility and Scotland) or in regions, such as Northern Italy, where early political disintegration strengthened kinship.⁷

These cultural changes in China and Europe took place on the eve of large population expansion and relocation. Climatic changes and invasions led to a large migration of Han people from the north to the east and south. These adverse conditions also weakened the state's ability to regulate migration (e.g., Kuhn, 2009, chapter 4). Large scale internal

⁵On China, see Fairbank (1994), Hamilton (1982), Khun (2009), Nee (2010). On Europe, see Goody (1983), Greif (2006, 2006a) cf. Herrmann-Pillath (2010).

⁶Confucianism emphasized intra-family moral obligations but not among common-descent relatives in general. Medieval noble clans kept genealogies.

⁷For survey and discussion, see Greif (2006), ch. 8 and Greif (2006a).

migration in the context of weak states also transpired in Europe as the population was recovering from its post-Roman decline.

Summarising, in both late medieval China and Europe large scale migration took place under weak states. How this occurred was influenced by initial cultural conditions: individuals could migrate and settle in the company of their kin, or with non-kin. But China and Europe had very different cultural conditions, for religious and political reasons unrelated to the phenomena that triggered the large scale migrations. As we shall see, these different initial conditions in turn had a profound and lasting influence over the resulting social arrangements.

3 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 whether individuals with given values settle among their kin (forming a clan), or in the company of non-kin (giving rise to a city). The role of the clan and the city is to sustain cooperation in the provision of a local public good. 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 organizational forms, and explores the joint dynamics of organizations and values. Throughout we abstract from the possible impact of the state, to which we return in section 4.3

3.1 Static Analysis

3.1.1 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 settle in the city or in their clan. There is a single city that can draw settlers from all dynasties, while the clan can only draw individuals belonging to a single dynasty; thus, there are several clans and one city. Settling 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 settled, 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 $H(\cdot)$ is continuously differentiable, concave, and strictly increasing,

with $H(0) = 0$ and $H_g(\tau M) > 1/M$ (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.⁸ Irrespective of where they choose to settle, 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 settle, 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. 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 ($\tau > \gamma$). Note that individuals can only contribute to the community which they joined (and which we metaphorically associated with a place of residence).⁹ Moreover, individuals choose where to settle, while nature chooses their type. Of course, as described below, identification (or morality) is one determinant of whether individuals choose to settle in the clan or in the city. For simplicity, the distribution of the idiosyncratic parameter δ is the same for clannish and generalist types.

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 3 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 3

⁸An equivalent formulation would have each individual bearing a cost p whenever it free rides or cheats on the community with which it identifies.

⁹Excluded by assumption is the possibility that one contributes to a faraway community or that a different members of a nuclear family live in different communities. This assumption is made for the ease of exposition.

	Location		
Type	<i>clan</i>	<i>city</i>	total
<i>clannish</i>	$x^{\lambda n}$	$x^{\lambda y}$	x^{λ}
<i>generalist</i>	$x^{\gamma n}$	$x^{\gamma y}$	x^{γ}
total	x^n	x^y	1

Thus, $x^{\lambda n}$ is the fraction of individuals who are of type λ (clannish) and settle in the clan, $x^{\lambda y}$ the fraction who is of type λ and lives in the city, and so on. The variables $x^p = x^{pn} + x^{py}$, for $p = \lambda, \gamma$, denote the proportion of type p in each dynasty and are chosen by nature. Similarly, $x^y = x^{\lambda y} + x^{\gamma y} = 1 - x^n$ denotes the fraction of each dynasty settling in the city, and it is determined in equilibrium, along with the proportions in each cell of Table 3.

With this notation and given the above assumptions, we can pin down the 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:

$$g^n = \tau x^{\lambda n} \quad (1)$$

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

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

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:

$$1/2 > e/M\tau \equiv \hat{e} \quad (3)$$

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

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

The timing of events is as follows: Each individual observes his type and then chooses whether to settle 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 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 settle 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, and 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.

3.1.2 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 settles 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 settle 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 settling 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 3.

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 by 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 $\delta^\lambda = \delta^\gamma + a$, 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 $p = \lambda, \gamma$ 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:

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

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 everyone 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 $Mx^{\lambda y}$ (the fraction of generalists settling 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.¹⁰

It is tempting to interpret these admissible equilibria as reflecting a historical transition that accompanies the growth in the size of cities. One 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 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 (values). 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).

¹⁰In 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.

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 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 $x^\lambda \in [x, \bar{x}]$, where $1 > \bar{x} > 1/2 > x > 0$*

(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>>

The appendix also proves that:

Proposition 3 *Suppose that the first inequality in condition (C1) holds. Then, in all three equilibria described in Proposition 1, city size is a strictly decreasing function of 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 settle in the clan while all the generalists settle 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) 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 2, to equilibrium (ii), to equilibrium (iii), city size increases. Propositions 2 and 3 thus imply 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.

3.2 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.

3.2.1 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 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.¹¹ 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):

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

¹¹Hu Hsien-Chin, the eminent scholar of China's descent groups noted that "families that move away ... retain their allegiance to the ancestral hall for many generations, the bonds of kinship being much closer than those of common residence" (1948, p. 10).

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:

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

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.

3.2.2 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 settling 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 settling in the city and all the clannish types settling 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 small 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 4 *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 settle in the clan while all the generalists settle in the city.

3.2.3 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 5 *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 settle 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 settling 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.

3.2.4 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 6 *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 settle in the city while the clannish mix between the clan and the city.

The specific conclusion that the majority of the population in this steady state settle in the city is due to the symmetrical structure of our model. Generally, in this steady state, more of the population settle in the city and less in the clan. It thus captures the social organizations that prevailed in Europe - hence the e superscript to denote this equilibrium.¹²

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.

3.3 Discussion

Contrasting the last three propositions, we see that even small differences in the initial social and moral conditions in otherwise identical societies lead to lasting and marked distinctions in social organization, morality, and institutions. Moreover, changes in the underpinnings of the clan and the city, by making each more or less attractive, have steady-state effects on citizens values. These permanent effects reflect the complementarity between morality and institutions in determining cooperation in the provision of 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, 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, benefiting from economies of scale in public good provision and the low cost of providing these goods based on limited morality. This situation perpetuates because the generalists living in the clan are more likely to have clannish offspring, compared to the generalists who settle in the city. We associate this equilibrium with the arrangements that prevailed in China, where loyalty to the clans that provided public goods was widespread.

A society that starts out with an emphasis on generalized morality finds itself in the opposite situation, captured by the equilibrium with segregation in the city. Here all the generalists remain in the city, while the clannish types are found in both locations. The city is large (and the clan is small) for two reasons. First, the city's large population, drawn from many dynasties, can provide more public goods. Second, since there are many generalists, free riding does not undermine the provision of public goods. The city is thus more attractive than the clan even for some clannish types. 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. We associate this equilibrium to the situation in Europe.

¹²For most of the period examined here, most Europeans lived in the country-side. These villages, however, were also characterized by being composed of individuals of distinct ancestry.

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 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 for why social institutions and morality evolved in different directions in these two parts of the world.

Interpreting the evidence in light of our analysis entails recognizing that the dichotomy between limited vs generalized morality is not absolute, but it epitomizes different metrics of moral obligations. Limited morality is often extended to individuals with kinship-like innate attributes such as place of birth or ethnicity. In contrast, generalized morality, although not always universal, it applies to individuals with acquired (as opposed to innate) characteristics, such as citizenship, class, residence, or religion. In either case morality, as anthropologists often noted, decreases with the distance from its defining characteristic.

Similarly metaphoric is the formal association of the city with an urban settlement and the clan with a rural one. Because the default in pre-modern societies was living among kin in a rural setting, generalized morality fostered urbanization. Yet, settlements composed of non-kin can also be rural, and cooperation amongst non-kin is based on generalized morality and formal institutions (as was common in Europe). Similarly a city can be composed of kin-based limited morality enclaves, as was common in China where such enclaves (rather than the city itself) were often in charge of local public good provisions. The main issue here is the overlap between the social organization toward which one feels moral obligation, and the unit providing local public goods.

4 History

Two societies that differ only in the initial distribution of values can theoretically evolve along different self-reinforcing trajectories of cultural traits, social organizations, and institutions. This section evaluates this result based on China's and Europe's historical experiences. Section 2 already documented the relative preminence of the city in Europe and the clan in China, as well as the origin of their different value systems. In this section we substantiate that each system evolved through positive feedbacks among social organization, morality, and institutions. The discussion transcends the boundaries of our model by incorporating the state in the analysis.

4.1 Enforcement and governance

Our analysis reveals why formal enforcement institutions are likely to be necessary in the city but not in the clan. Historically, in fact, cooperation in the European cities and the Chinese clans was sustained by different methods.

European cities increasingly developed formal, legal institutions to support cooperation and there was a transition from legal customs to legal codes and from elected voluntary judges

to professional ones (e.g. Clark, 1987). Clearly, cooperation was also based on private-order institutions. Among these were organizations (such as guilds) that provided club-goods and fostered cooperation among non-kin by the threat of exclusion.

Although clans formalized their internal regulations in rule-books and had adjudication procedures, intra-clan legal disputes were arbitrated by the elders and compromise was the goal. Moreover, the law was based on customs and not legal principles. In terms of our analysis, limited morality reduced intra-clan enforcement cost, thus reducing the potential gains from creating more formal legal institutions.¹³

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 Song Dynasty (960-1279) 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).

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. This implies that intra-city public good provisions should have been provided by either the state or by each clan to its members. Indeed, Chinese cities were governed by state-appointed officials and residential areas were divided into smaller wards, called fangs, which, served as neighborhood administrative units. Each fang was usually dominated by people from the same clan or region (Rowe, 1984; Huang, 2006). Prior to the Song Dynasty (960-1279), the fangs were surrounded with walls and guarded gates (e.g., Huang, 2006). This urban structure reveals and reinforces limited morality.

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 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 economic interest, became common only in the 19th century. Only 268 interest-based guilds are known to have operated in China from 1655 to 1911. Most of them were founded in the 19th century (Moll-Murata, 2008).

4.2 Institutions, exchange, and morality

The institutional foundations of markets in Europe and China also differed as predicted by distinct moralities which they arguably positively reinforced. In Europe, intra-city formal enforcement institutions supported inter-city impersonal exchange through the Community

¹³See review in Ma (2007). Cf. Zelin et al (2004).

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, 2006a). Trade would cease following an uncompensated default, inducing the city to compensate for the default and punish the one who defaulted to deter others. General morality renders this punishment more credible.

Under limited morality a CRS is less effective. To see why, note that *only* under generalized morality both intra-city and inter-city impersonal exchange relies on legal contracting. Under limited morality, exchange with a clan member – even from afar and without personal 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 diasporas that managed 18th century interregional trade usually were built upon kinship ties. The huge shipments of rice ... were overseen by groups of Kiangsi merchants organized internally by lineage” (Rowe, 2002, pp. 531-2; cf Ma, 2004).¹⁴

It stands to reason that the implied patterns of interactions fostered their underpinning moralities. In Europe, the Community Responsibility System implied equality before the law, thereby reinforcing this moral principle. It also implied a shared interest in promoting generalized morality because outsiders held every city member to be jointly liable for misconduct by others, thereby providing incentives. Equal treatment of out-of-town traders was generally mandated in European cities’ trade regulations and 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 commercial networks. This positive feedback to avoid inter-clan relations is reflected in clans’ rule books from the early 20th century. Although “friendship [as distinct from kin-relations] 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).

4.3 The scope of morality

Whether charity is personal or impersonal is a good measure of moral obligation toward non-kin. Charity is personal when the giver donates to specific individuals he knows and it is impersonal otherwise. If limited morality prevailed in China, personal charity among kin was more likely to predominate. In contrast, if generalized morality prevailed in Europe, impersonal charity to non-kin was more likely to predominate. This, indeed, was the case.

Charity in pre-modern China was generally given to kin. The innovator of the clan trust, 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). Only in the early 17th century non-Buddhist, impersonal charity permanent organizations were established

¹⁴For morality in overseas Chinese trading networks see, e.g., Landa (1981), Pyatt and Redding (2000), and Redding (1993).

on some scale.¹⁵ Although the Chinese authorities encouraged impersonal charity, moral philosophers decried it viewing the diversion of assistance away from kin immoral. A popular 17th century morality book “tells of a generous scholar who was derided by a member of his lineage for lightly giving money away to strangers” (ibid).

In contrast, by the 16th century impersonal charity was long established in Europe. To illustrate, in 1560, there were about 35 poor relief-foundations in London and their number increased to more than 100 by 1700. Donations for the "relief of the poor" or the "poor man box" appear in a quarter of the wills in late 16th century Bristol and a century later this fraction had been about 30 percent in England as a whole (Ben-Amos, 2000).

The low level of impersonal charity in China manifested itself in wife selling. Very poor families – arguably those without a supportive kin group – were dissolved by the selling of the wife (with or without the children) to another man to become his wife or concubine (Sommer, 2007). 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 when unable to support her. From 1750 to 1900 poverty motivated more than 75 percent of the sales (Sommer, 2007). In Europe the practice was very rare arguably because the very poor had access to poor relief provided by non-kin.

The morality of a pre-modern society can also be extrapolated from its criminal law, because 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. European legal codes exhibit uniformity and emphasized equality before the law. In European cities everyone was *de jure* equal under the positive (man-made) laws and in European states even rulers were under the ‘normative’ (divine) law.

In contrast, under Chinese law, punishments 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).

Legal diversity prevailed 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, p. 164, fn 37). Moreover, 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).

¹⁵Smith (1987) summarizes: “during the late Ming and early Ch'ing (1580- 1750), ...[new organizations were devoted] to voluntary or compassionate giving to the poor and needy outside one's family” (pp. 319-21).

4.4 Society and State

Although the state is not captured in our formal analysis, this subsection highlights a positive political feedback that reinforced the importance of the clan and of the city in China and Europe respectively.

Pre-modern European rulers extensively relied on cities' administrative capacity in pursuing their interests confirming role as units of cooperation. Cities collected tax, provided navies, mustered armies, and administered justice. Intra-city enforcement institutions and generalized morality facilitated intra-city, inter-lineage cooperation in confronting feudal lords and mitigating internal conflicts. Cities' administrative power vis-à-vis the state (Greif, 2005) enabled most cities in Western Europe to gain self-governance by 1500.

The eventual transition in European states to more centralized administrative and legal structures was arguably facilitated by the prevalence of generalized morality and the associated legal infrastructure and principles. In England, for example, the merchant law was integrated in the Common Law and in France, Code Napoleon built on Paris' legal code.

The lack of self-governed cities in China was not simply due to a more powerful state, but also to a 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" (Friedmann, 2007, p. 274). As noted above, families that moved to cities retained "their allegiance to the ancestral hall for many generations, the bonds of kinship being much closer than those of common residence" (Hsien-Chin, 1948, p. 10).

The importance of the clan as a unit of cooperation in China is evident from their administrative roles. "Because of the feeling of mutual responsibilities and because of its ability to maintain high moral standard among its members... the state has found it convenient to leave to it [the *tsu* (clan) organization] considerable judicial powers" (ibid). In terms of our analysis, relying on this quasi-private enforcement was preferred by the central authorities despite the economies of scale in law enforcement, because clan enforcement took advantage of the disciplinary impact of limited morality and reputation.

In China, "a localized lineage was normally the largest unit within which disputes between its members were settled" (Freedman, 1966, p. 114). Clans were expected to specify and enforce the rules governing the relations among their members. Similarly, clan rules regularly discouraged litigation and favored arbitration provided by the clan (Liu, 1959). The Chinese state thus 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 was less required and would have undermined the clans, an outcome contradictory to the interests of the elders who controlled the clans and the state that used them. Reliance on intra-clan enforcement explains why China did not develop commercial legal code until the late 19th century and why compromise was the objective of civil adjudication.¹⁶ At the same time, limited morality and a tradition of private adjudication and unequal legal rights, however, have arguably reduced the gain from investing in legal infrastructure.

¹⁶For the debate about China's legal development see Nakamura (2004) and Ma (2007).

Similarly, the state also delegated administrative authority to kinship groups either directly or indirectly by relying on local leaders (e.g., Huang, 1985. pp. 224-41; Heijdra, 1998, pp. 468-70). Reducing the size of the government by relying on clans was an explicit policy goal of Chen Hongmou (1696-1771) who was regarded, both in his day and after, as a model provincial governor. He granted lineage headmen considerable judicial and disciplinary powers over their kinsmen in 1742 (Rowe, 1998, p. 378). Similarly, the reformer Gu Yanwu (1613-82) argued that increasing the lineage's judicial authority would reduce the judicial workload and hence the requisite size of the bureaucracy (ibid, p. 383).

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 number of public officials remained almost the same from the 10th to the 19th century despite a large increase in population and territory (ibid).

Because Chinese clans economized on enforcement cost, the state created complementary institutions. It promoted the ideology of filial piety, the moral obligation toward parents and ancestors. Although filial piety had been a cornerstone of Chinese ethics since the Han Dynasty (206 BCE– 220 CE), the Song (960-1279) developed, codified, and popularized it (Hamilton, 1984). 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 the conduct of their members. In general, laws backed the power of the kinship organizations – and the elders who controlled them – over their members.

In particular, the Chinese state reinforced the clan – and the social control it enables – through the *hukou* (household registration) system that limited mobility. The effectiveness of this system notwithstanding, it does not seem to have been the reason for the clan-city dichotomy. State control was weak particularly when natural and man-made disasters induced migration.

The power of clans could become threatening for the State rulers, as evident in the clannish basis of China's dynasties (e.g., Song, Ming, Qing, etc.). These ruling clans were well aware of the political risk that large kinship groups implied. "The state frequently saw clans as a potential source of disputes and public disturbances" (Ruskola, 2000, pp. 1662-3). At the same time, the alternative of strengthening the provincial administrations was also politically risky. To mitigate the adverse affect of relying on clans, the state took such measures as avoiding forming local militia to the extent possible and fixing the land tax level in the 18th century.

More importantly, the examination system mitigated the risk from large clans by aligning their interests with the state's. Powerful clans had the human and financial resources required to field successful candidates who, upon becoming public officials, benefited their clans.¹⁷ A public official and his descendents had a political voice and legal privileges that were bestowed, to some degree, on his relatives. These privileges lasted for three generations, therefore insuring that previously powerful clans would not for ever crowd out rising ones.

In Europe, cities, and not clans, threatened the power of the State. To cope with this threat, Europan cities (like Chinese clans) were involved in the governance of the State.

¹⁷Although the examination system predated the clan system, they were institutionalized during the Song.

Although cities were prone to revolt, self-governed cities remained integral part of states' administrative apparatus even in strong states such as France or Germany after its unification in 1871. In general, cities had political voice, and privileges were allocated on a territorial and quasi-permanent basis by incorporating large and prosperous cities in representative assemblies. Cities were also directly represented in the twenty-five national and provisional assemblies that existed in Europe by the end of the 15th century.

5 Cultural Persistence and the Reemergence of Clans in Modern China

The 20th century provides additional evidence supporting the mutual reinforcement between morality and social organizations. This section documents the persistence of the cultural attitudes described in the previous section, and the re-emergence of clans in modern China.

5.1 Cultural persistence

We start by documenting how cultural attitudes remain different today between modern China and Europe, focusing in particular on Chinese immigrants in the West.

In contemporary China limited morality is reflected in trust attitudes. In 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).¹⁸ The 2006 China General Social Survey asked villagers how much they trust a villager with the same surname vs a villager with another surname (qf14). An unconditional mean comparison test of the resulting 4228 responses in random sample of 76 counties reveals a statistically significant lower trust (at the 1 percent level) in villagers of different surname.¹⁹

The World Value Surveys allow one to compare attitudes across countries. In the first time China was included (1990), 60 percent of the Chinese responded that most people can be trusted.²⁰ This is much higher rate than in the West (i.e., France, GB, USA, and Germany) led by the USA with 51.2 percent. The response to a more specific question about trust, however, reveals the opposite. Only 11.6 percent of Chinese "trust completely" other "people in the country" compared to an average of 16 percent in the West. Moreover, the World Value Survey of 1990 is the only one to ask about trust toward individuals of different social distance. The respondents were asked to “say how much you trust the following groups of people.” The level of trust in family was the highest in both China and the US. The levels of trust toward strangers, however, differed. In the US, they remained relatively high, about 50 percent, while in China they dropped to 10 percent (Buchana and Crosonb, 2004, pp. 489-490). Later surveys conducted in 2006 and 2007 asked about trusting a person whom

¹⁸Chinese overseas businessmen perceive Westerners as more contractually reliable. A survey of Chinese businessmen in Thailand and Hong-Kong (1994-5) 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).

¹⁹See on-line appendix on our web pages.

²⁰WVS, Roland Inglehart, et al 1998 and the World Value Survey, 2005-8. www.worldvaluessurvey.org.

one met for the first time. 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. While almost 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. Complete trust in people that one knows personally is 20.5 percent in China, but 31.6 percent in the West.

To overcome the difficulty of comparing attitudes across countries, where the same question can have different interpretations, we collected data on the attitudes of Chinese vs European immigrants to Canada, exploiting the 2005 Ethnic Diversity Survey. We restrict attention to immigrants of second generation or higher. These are individuals born in Canada, and who report their ethnic ancestry.

We focus on trust towards different reference groups, exploiting a question that asks "How much do you trust each of the following groups: people in your family / people that you work with / people in your neighborhood?". The answer can vary from 1 (not at all) to 5 (a lot). We also compare trust attitudes in general, with the usual question on trust towards people in general ("Generally speaking would you say that most people can be trusted or that you cannot be too careful in dealing with people?", with a binary answer (1 if people can be trusted, 0 otherwise). Compared to European immigrants, we expect Chinese immigrants to display more trust in people who are closer (such as family members), and less trust in people who are more distant (such as in the general trust question). More generally, as familiarity or closeness decreases, we expect trust to decline more steeply for Chinese immigrants compared to European immigrants.

Table 4 reports the estimates for the different trust attitudes. Estimation is by OLS for trust in family members, colleagues and neighbors, and by Probit for general trusts (marginal effects are reported). Thus, the coefficient estimates are strictly comparable only for the first three measures of trust, and not for general trust. The dummy variable *Chinese vs European* takes a value of 1 for Chinese ancestry, and 0 for ancestry in Northern and Central Europe. It is coded as missing for ancestry from other parts of the world.²¹ Throughout we also control for individual education, age, income, gender, and other observable individual features (see the Notes to Table 4). As shown in Panel A of Table 4, Chinese and European immigrants have similar levels of trust towards family members and colleagues at work (columns 1 and 2). When it comes to trust towards neighbors or people in general, however, Chinese immigrants are significantly less trusting than European immigrants (columns 3 and 4).

Could these results reflect the possibility that Chinese immigrants find it more difficult to fully integrate in a Western society, because of their different ethnic traits? To address this concern, Panel B of Table 4 adds a regressor that controls for whether the respondent ever felt uncomfortable up until the age of 15 because of his /her ethnic identity (the variable *uneasy*). With this specification, the results are unchanged or even stronger. Chinese immigrants have relatively more trust in family members compared to European immigrants. This difference remains positive but it loses significance with respect to colleagues, and it becomes negative and statistically significant when turning to neighbors and to general trust.

Insert Table 4 here

²¹Thus, immigrants from Southern Europe (Italy, Spain, Greece and Portugal) are excluded from the comparison, because trust is traditionally low in this part of Europe- indeed the term *amoral familism* was introduced by Banfield (1967) precisely in his analysis of Southern Italy.

Altogether, these results confirm that the different cultural traditions of China and the West continue to shape attitudes today.

5.2 The Persistence and Reemergence of Clans

The persistence of cultural attitudes is matched by a striking persistence of clans as a central organization in Modern China.

The modernization movement in the early 20th century was hostile to the clans, that were viewed as an obstacle to economic development. In 1904, the Chinese government legalized corporations with the explicit intention to foster joint stock companies. The law failed in this regard, as Li Chun explains: "the idea that members of the public would be invited to join one's business and share in its control and profits was indeed repugnant. On the other hand, the notion that one's money be put into the pocket of some strangers for them to run a business was just as unthinkable" (Li, 1974, p. 205 cited by Kirby, 1995, p. 50).

The communist regime officially abolished the clans upon 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., Huang, 1985, p. 308). Had the clans been a product of the state, they would not have survived the crackdown since 1949. If, however, clans had been a product of the coevolution of deeply held moral convictions, social organization, and institutions, clans should have persisted and reemerged following the reforms that allowed individuals to organize themselves. This is indeed what has happened since 1978.

A county-level survey in 2000 (by Liangqun and Murphy, 2006, in Jiangxi) documents that 70 surnames out of 99 (in 40 villages) updated their genealogies since 1981 and 41 surnames invested in their ancestral shrines since 1991 (p. 230). A 2002 representative national survey of more than 300 villages reveals lineage activities and kinship organizations in 66 percent of the villages (Tsai, 2007, pp. 154–7). Clans resumed their role in securing property rights from predation by officials, organizing weddings and funerals, providing welfare, contributing to public projects, and promoting mutual aid arrangements (ibid). Inter-clan conflicts also resumed and collectively owned rural firms often formally exclude non-locals (Thøgersen, 2002). About 90 percent of the 887 households that migrated to or from one of 50 villages relocated to their 'ancestral village' and 60 percent relocated due to inter-lineage tension (Liangqun and Murphy, 2006, p. 623).

We quantify clans' persistence using a random sample of 76 counties, 205 villages and 4274 individuals from *China General Social Survey, 2005* (GSS05). The GSS05 asks (only) rural residents whether there is a clan organization in their community and, whether it is a surname-based or a temple-based *clan organization*.²² Although under-reporting of clan organizations is likely given tradition of suppression by the communist authorities, the census reveals 277 clan-based organizations. A clan organization almost always (90 percent) has a genealogy, a graveyard, or both. The two organizational types differ, however, in their economically-relevant assets such as land, estates (other than ancestral hall), and trust funds. Only 26 percent of the surname-based network have such an asset compared to 78 percent of

²²F10. '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. '

the temple-based organizations (F12). On average there is 1.35 organizations per-village and one organization per 15.5 respondents. The highest number of people per-clan organization is in the eastern region (35) and the lowest is the northwest (8).

Almost 70 percent of the population live in a county with positive sample probability of a village having an organization and in 41 percent of the counties the village-probability of having a clan organization is at least 50 percent. In fact, clan organizations currently exist in each of China's six regions although there are no temple-based clan organizations in the north (Figure 2, note that the northwest is a separate region) but they are particularly strong in the south-central region and, specifically, in Guandong, the richest province. These findings correspond to our historical discussion in which we noted the relations between out-migration from the North and the rise of clans.

<<insert figure 2 around here >>

Table 5 presents the number and fraction of organizations – out of the 277 in our sample – that fulfill 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 5 around here >>

Our on-line appendix also documents that the frequency and strength of clans in modern China is negatively correlated with urbanization in townships (i.e. excluding major cities), as predicted by the model. Specifically, data from China County-prefectural Statistical Yearbook, 2006 (YB06) and China General Social Survey, 2005 (GSS05) reveals that in the sample of 76 counties, clan strength is negatively correlated with urbanization even after controlling for such variable as education, infrastructure investment, and distance from the coast.

This re-emergence of clans is particularly noteworthy given that the reforms were not designed to foster clan organizations.²³ Households, and not clans, were given land-use rights in the former collective farms and privately-owned businesses were permitted. Yet, kin-based and relations-based exclusive organizations have re-emerged and resumed their traditional role in supporting cooperation.

5.3 Clannishness and Market Reforms

Chinese scholars have noted the important role of kin and kin-like relations in the success of the market reform. Clans secure property rights from local officials, organize weddings and funerals, provide welfare, contribute to public projects (e.g., road construction), and promote 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 are significantly correlated with the provision of public goods and with

²³Yet 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., Huang, 1985).

how public officials are held accountable (Tsai, 2007). There is also a strong and significant correlation of village-level kinship with the number of private enterprises and the labor force (Peng, 2004).²⁴

More generally, the reform led to proliferation of institutions based on limited, relations-specific morality and affinity such as *guanxi* (relations; e.g., Thøgersen, 2002, p. 261) and clan-like organizations.²⁵ *Guanxi* reflects and reinforces limited morality, also because of the services it provides to individuals in their economic activity. To illustrate, job assignment in the 1980s was formally based on meritocracy. Yet, in a representative sample of 1,008 households in the city of Tianjin (1988), 57 percent reported that they got their jobs (1977–1988) based on personal connections (*guanxi*) and 43.3 percent of those who assisted them were relatives (Bian, 1994).

China's private-order institutions – based on limited morality and reputation among family members, friends, and business associates – have contributed to the success of the market reforms. Indeed, they seem to be an important reason why the reforms succeeded though China's formal institutions do not adequately secure private property rights or enforce contracts.²⁶ The deficiency of China's formal institutions, as compared to Europe's, is well known. Prior to 2007 all property in China was de jure owned by the state and there was no law recognizing and regulating private property rights.²⁷ There was no dedicated contract law in Communist China prior to 1981 and only in 1999, a Uniform Contract Law had been formulated. This law, which is still in force, is ineffective because the legal system is subservient to the interest of the state and "it is not only hard to find the law but also difficult to understand them... [and it is] hard to use the law in practice because of its famous ambiguous" (Li, Grace, 2009, p. 20). The inadequacy of China's formal institutions manifests itself in the widespread reliance on political actors (e.g., party cadres) in business enterprises (e.g., Nee, 1992) and the evidence suggests that 'buying' political protection in this manner is costly.²⁸

The cultural and institutional foundations of China's contemporary economy have much in common with those that prevailed in Imperial China. Yet, they differ in important ways. In particular, the return to market economy in China has transpired in an institutional context in which clans are no longer economic units with legal power. The interest and resources of clans' leaders no longer stand in the way of labor mobility, risk-taking, and interacting with non-kin. Limited morality fosters individual-level profit-seeking market behavior on the one hand and relations-based institutions for collective action and contract enforcement on the other. Altogether, these results illustrate the persistence of the clan and their impact on modern China.

²⁴Data from 366 villages in 22 counties. The effect is smaller including self-employed but significant; 7 percent increase per 10 percent increase in the proportion of the largest group.

²⁵Regarding fictional clans in Late Imperial China, see Rowe (2002), pp. 355-7.

²⁶E.g., Nee (1992) and Whyte (1995, 1996).

²⁷E.g., Zhang (2008); Yet, article 11 of the 1988 constitutional amendment permitted private enterprises.

²⁸E.g., Bai et al (2006), Lu and Tao (2010), Su and He (2010), Li et al (2008), Lu et al (2010), Thøgersen (2002).

6 Differences within Europe

This section exploits heterogeneity within Europe and presents empirical evidence that urbanization was less diffuse in European areas where kin-based relations were stronger.

6.1 Urbanization in the European regions

Kin also played an important role in Europe, particularly in providing safety nets, agency service in long-distance trade, and protection (e.g., Ben Amos, 2008; King, 2000, p. 215). As a measure of the strength of kin-based relations, we rely on the family structure. Todd (1990) 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 much smaller scale (e.g., Wolf, 1955 and Bresc, 1996). 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 proxy for the strength of kin ties in various European regions, and more generally 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, 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). 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 classification of family structures used in the analysis captures remarkably stable traditions dating to the Middle ages if not earlier (cf. Todd, 1990). Nevertheless, the theory also predicts that urbanization in turn favors the diffusion of the nuclear family and a gradual erosion of the extended family tradition. Thus, the possibility of reverse causation cannot be ruled out, and the regressions below can be interpreted as correlations consistent with the theory, not as causal relations going from family structure to urbanization.

6.1.1 Data

Our data is built on Duranton and Rodriguez-Pose (2009), 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 (2009), regions are defined at the NUTS II level, according to the 2003 Eurostat classification. However, in order to keep data consistent 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 Chèvre (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).²⁹ We have thus imputed population in 1600 to each region of these countries under the assumption that the share 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, although it obviously contains some measurement error.

The distribution of *urbanization around 1600* in our sample is illustrated in Figure 3a. It has a mean of 8.8 percent and it ranges from 0 (for several regions) to 51 percent, for Andalusia.

<<Insert Figure 3a, about here>>

Family types Our classification of family types is based on the work of Todd (1990), and 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 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 (1990) measures the diffusion of both family types across Western Europe. He combines contemporaneous information (having removed the effect of rural employment) 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.³⁰

²⁹Their 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.

³⁰Todd (1990)'s classification is finer than the one we use. He also classifies families according to whether they treat children equally but this distinction is not relevant for our purposes.

Duranton and Rodriguez-Pose (2009) 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 (2009) 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 (2009).

Some areas are classified by Todd (1990) as being dominated by no particular type of family, or as not 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 3b. It has a mean of 60.7 percent, and ranges from 0 to 100 percent. Contrasting the two maps in Figure 3 suggests that regions where the extended family prevails (darker regions) seem to be associated with a lower rate of urbanization (lighter regions), as predicted by the theory.

<<Insert Figure 3b, about here>>

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 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.

6.1.2 Results

Table 6 presents the results. The dependent variable is always *urbanization around 1600*. 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 6 around here>>

Figure 4 illustrates the correlation in the data (after removing the effect of all regressors included in the least parsimonious specification of column 4 in Table 6). Several outlier observations are present, but none of them individually is driving the correlation.

<<Insert Figure 4 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.

7 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 moralities in multiple ways that our parsimonious model does not capture. These effects persist today.

Persistence, however, does not mean stasis and both systems 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.³¹ For example, the Europeans responded to the technologically driven increase in economies of scale by relying on the law and general morality to

³¹For this view of culture see, for example, Greif (1994, 2006) and Herrmann-Pillath (2010). Regarding China, see Thøgersen (2002); Herrmann-Pillath (2009); Redding and Witt (2009).

separate between ownership and control. 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 has a comparative advantage in economic activities in which personal relations are more efficient. The relative efficiency of the Chinese and European systems is thus theoretically unclear, as each system has advantages and drawbacks. 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. However, 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. Yet, limited morality and informal institutions foster potentially destructive inequality, corruption, and social tensions.

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 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.

8 Appendix

8.0.3 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.

8.0.4 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):

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

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):

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

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

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

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

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

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

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:

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

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

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

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

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

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:

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

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:

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

Clearly, $x^{\max} < 1 - \hat{e}$. 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) $D^c(x^{\max}) > 0$, we must also have $x^{\max} > \bar{x}$. Hence if $x^\lambda \in (\bar{x}, x^{\max})$ 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:

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

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

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

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 $x^\lambda \in (\bar{x}, x^{\max})$, where $1 - \hat{e} > x^{\max} > \bar{x} > 1/2$. 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:

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

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

$$x^{\lambda n} = d\delta^\lambda x^\lambda \quad (23)$$

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

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

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:

$$1/d > D^e(x^\lambda) > 0 \quad (25)$$

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 (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:

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

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

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

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 $x^y \in ((1-\underline{x}), 1)$.

8.0.5 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:

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

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:

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

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:

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

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. 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:

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

QED.

8.0.6 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:

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

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

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

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 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:

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

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.

8.1 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 Comunidad 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 latter is geographically distinct. 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 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 (2009) 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 have been kindly provided by Jan Luiten Van Zanden.

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>, in turn relying on the following sources: Almanach de Gotha, census: Italy (1861), Netherlands (1859/60), Spain (1860), U.K. (1861), Ireland (1861), Sweden (1860/1861), Western Germany (1867, estimate), Portugal (1862, estimate), Finland (n.s., 1861). French Statistical Institute INSEE, census: France (1861). Statistisches Jahrbuch fur Osterreich, census: Austria (1857).

9 References

- Andersson P.L. 2000. "Folkmängden i Sveriges socknar och kommuner 1571-1991"
- Bairoch, P., J. Batou and P. Chèvre. 1988. *La population des villes Européennes de 800 á 1850*. Genève: Libraire Droz.
- Bai, C.E., J. Lu, and Z. Tao. 2006. "Property Rights Protection and Access to Bank Loans: Evidence from Private Enterprises in China." *Economics of Transition* 14 (4): 611–628.
- Banfield , E.C., 1967. *The Moral basis of a Backward Society*. New York, NY: Free Press.
- Ben-Amos, I.K. 2000. "Gifts and Favors: Informal Support in Early Modern England." *The Journal of Modern History*, 72(2): 295-338.
- Ben-Amos, I.K. 2008. *The Culture of Giving: Informal Support and Gift-Exchange in Early Modern England*. Cambridge, UK: Cambridge University Press.
- Beresford, M. and H. P. R. Finberg. 1973. *English Medieval Boroughs: A Handlist*. Newton Abbott: David and Charles.
- Bian, Y. 1994. "*Guanxi* and the Allocation of Urban Jobs in China." *The China Quarterly*, 140: 971-999.
- Bisin, Alberto & Verdier, Thierry, 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory*, vol. 97(2): 298-319.
- Boerner, L. and D. Quint. 2010. "Medieval Matching Markets." Memo. Free University Berlin.
- Brandt, L., D. Ma, and T.G. Rawski. 2011. "From Divergence to Convergence: Re-evaluating the History Behind China's Economic Boom." WP. University of Toronto.
- Brenner, R. and C. 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.
- Bresc, H. 1996. "Europe: Town and Country (Thirteenth-Fifteenth Century)" in André Burguière et al. (ed) *A History of the Family*, Vol. I Cambridge: Polity Press. 430-66
- Broadberry S., B. Campbell and B. van Leeuwen. 2010. "English medieval Population: reconciling time series and cross sectional evidence", working paper.
- Buchana, N. and R. Crosonb. 2004. "The boundaries of trust: own and others' actions in the US and China." *Journal of Economic Behavior & Organization*, 55: 485–504.
- Clark, D. S. 1987. "The Medieval Origins of Modern Legal Education: Between Church and State." *The American Journal of Comparative Law*, 35 (4): 653-719.

- Deng, K.G. 2000. "A Critical Survey of Recent Research in Chinese Economic History." *Economic History Review*, LIII: 1-28.
- Duranton, G. and A. Rodriguez-Pose 2009. "Family types and the persistence of regional disparities in Europe." *Economic Geography*, 85: 23-47.
- Fairbank, J.K. 1994. *China, A New History*. Cambridge, MA: Harvard University Press.
- Fei, J. C. H. and T.J. Liu. 1982. "The Growth and Decline of Chinese Family Clans." *Journal of Interdisciplinary History*, 12(3): 375-408.
- Freedman, M. 1966. *Chinese Lineage and Society: Fukien and Kwangtung*. London: Athlone.
- Friedmann, J. 2007. "Reflections on Place and Place-making in the Cities of China." *International Journal of Urban and Regional Research*, 31(2): 257-279
- Gellhorn, W. 1987. "China's Quest for Legal Modernity." *Journal of Chinese Law*, 1(1): 1-22.
- Goody, J. 1983. *The Development of the Family and Marriage in Europe*. Cambridge, UK: Cambridge University Press.
- Greif, A. 1994. "Cultural Beliefs and the Organization of Society." *The Journal of Political Economy*, 102(5): 912-950.
- Greif, A. 2005. "Commitment, Coercion, and Markets." in C. Menard and MM. Shirley (eds), *Handbook for New Institutional Economics*. New York, NY: Springer.
- Greif, A. 2006. *Institutions and the Path to the Modern Economy*. Cambridge, UK: Cambridge University Press.
- Greif, A. 2006a. "Family Structure, Institutions, and Growth." *American Economic Review*. 96(2): 308-312.
- Greif, A. 2012. "A Theory of Moral Authority," WP. Stanford University.
- Guiso, L., P. Sapienza and Luigi Zingales. 2006. "Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives* 20: 23-48.
- Hamilton, G.G. 1984. "Patriarchalism, in Imperial China and Western Europe: A Revision of Weber's Sociology of Domination." *Theory and Society*, 13(3): 393- 425.
- Herrmann-Pillath, C. 2009. "Social Capital, Chinese Style: Individualism, Relational Collectivism and the Cultural Embeddedness of the Institutions-performance Link." *China Economic Journal*, 2(3): 325-350.
- Herrmann-Pillath, C. 2010. "What Have We Learnt from 20 Years of Economic Research into Culture?" *International Journal of Cultural Studies*, 13(4): 317-335.

- Heijdra, M. 1998 [2007], "The Socio-economic Development of Rural China During the Ming." In D. Twitchett and J. K. Fairbank (eds) *The Cambridge History of China*, vol. 8(2). Cambridge, UK: Cambridge University Press. Ch. 9, 417-581
- Hsien-Chin, H.. 1948. , *The Common Descent Group in China and its Functions*, New York: Viking Fund, Viking Fund publications in anthropology, no. 10.
- Huang, P.C. 1985. *The Peasant Economy and Social Change in North China*. Stanford, CA: Stanford University Press.
- Huang, Y. 2006. "Collectivism, Political Control, and Gating in Chinese Cities." *Urban Geography*, 27(6): 507–525.
- Inglehart, R., M. Basanez, and A. Moreno. 1998. *Human Values and Beliefs. A Cross-Cultural Sourcebook*. Ann Harbor: Michigan University Press.
- Kirby, W.C. 1995. "China Unincorporated: Company Law and Business Enterprise in Twentieth-Century China," *The Journal of Asian Studies*, 54(1): 43-63.
- King, S. 2000. *Poverty and Welfare in England 1700-1850*. Manchester, UK: Manchester University Press.
- Kuhn, D. 2009. *The Age of Confucian Rule. The Song Transformation of China*. Cambridge, MA: Harvard University Press.
- 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.
- Li, C. 1974. "The Kung-ssu-lii of 1904 and the Modernization of Chinese Company Law." *Zhengda faxue pinglun (Chengchi University Legal Review)*, 11: 163-209.
- Li, G. 2009. "The PRC Contract Law and Its Unique Notion of Subrogation." *Journal of International Commercial Law and Technology*, 4(1): 12-21.
- Li, H., L. Meng c, Q. Wang d, L.A. Zhou. 2008. "Political Connections, Financing and Firm Performance: Evidence from Chinese Private Firms." *Journal of Development Economics*, 87(2): 283–299.
- Liangqun, L. and R.l Murphy. 2006. "Lineage Networks, Land Conflicts and Rural Migration in Late Socialist China." *Journal of Peasant Studies*, 33(4): 612-645.
- Liu, H.C.W. 1959. "The Traditional Chinese Clan Rules." *Monographs of the Association for Asian Studies*, 7.
- Lu, J. and Z. Tao. 2010. "Determinants of Entrepreneurial Activities in China." *Journal of Business Venturing*, 25(3): 261 -273.
- Lu, J., Z.Tao, and Z. Yang. 2010. "The Costs and Benefits of Government Control: Evidence from China's Collectively-owned Enterprises." *China Economic Review*, 21(2): 282 -292.

- Ma, D. 2004. "Growth, Institutions and Knowledge: a Review and Reflection on 18th-20th Century Chinese Historiography." *Australian Economic History Review*, 44(3): 259-77.
- Ma, D. 2007. "Law and Economic Growth: the Case of Traditional China- a Review with Some Preliminary Hypotheses." Working Paper. London School of Economics.
- McEvedy C. and Jones R. 1978. *Atlas of World Population History*. New York, NY: Penguin Books.
- McKnight, B.E. 1985. "Song Legal Privileges." *Journal of the American Oriental Society*, 105(1): 95-106.
- Mocarelli, L. 2008. "Guilds Reappraised: Italy in the Early Modern Period." *International Review of Social History*, 53: 159–178.
- Moll-Murata, C. 2008. "Chinese Guilds from the Seventeenth to the Twentieth Centuries: An Overview." *International Review of Social History*, 53: 213–247.
- Nadal, J.1984. "La poblaciòn española (siglos XVI a XX)" Ariel.
- Nakamura, S. 2004. "Was Traditional Chinese Law a Mere Model? Part Two." *International Journal of Asian Studies*, 1(2): 297-322.
- Nee, V. 1992. "Organizational Dynamics of Market Transition: Hybrid Forms, Property Rights, and Mixed Economy in China." *Administrative Science Quarterly*, 37(1): 1-27.
- Nee, V. 2010. "Bottom-up Economic Development and the Role of the State." Cornell University Working Paper #49.
- Pirenne, H. 1969. *Medieval Cities: Their Origins and the Revival of Trade*. Princeton, NJ: Princeton University Press.
- Peng, Y. 2004. "Kinship Networks and Entrepreneurs in China's Transitional Economy." *American Journal of Sociology*, 109 (5): 1045–1074.
- Pomeranz, K. 2000. *The Great Divergence: China, Europe and the Making of the Modern World Economy*. 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): 41-63.
- Redding, G. and M.A. Witt. 2009. "China's Business System and its Future Trajectory." *Asia Pacific Journal of Management*, 26: 381-399.
- Redding, S. G. 1993. *The Spirit of Capitalism*. Berlin: Gruyter.
- Rosenthal J.L. and R.B. Wong. 2011. *Before and Beyond Divergence. The Politics of Economic Change in China and Europe*. Cambridge, MA: Harvard University Press.
- Rowe, W.T. 1984. *Hankow: Commerce and Society in A Chinese City, 1796-1889*. Stanford, CA: Stanford University Press.

- Rowe, W.T. 1998. "Ancestral rites and political authority in late imperial China - Chen Hongmou in Jiangxi." *Modern China*, 24(4): 378 -407.
- Rowe, W.T. 2002. "Social Stability and Social Change." In D. Twitchett and J.K. Fairbank (eds), *The Cambridge History of China*, vol. 9. Cambridge, UK: Cambridge University Press. Ch. 9, 473-562.
- Ruskola, T. 2000. "Conceptualizing Corporations and Kinship: Comparative Law and Development Theory in a Chinese Perspective." *Stanford Law Review*; 52(6): 1599-1729.
- Skinner, W.G. 2001 [1974]. *Marketing and Social Structure in Rural China*. Ann Arbor, MI: The Association of Asian Studies.
- Smith, J.F.H. 1987. "Benevolent Societies: The Reshaping of Charity During the Late Ming and Early Ch'ing." *The Journal of Asian Studies*, 46(2): 309-337.
- Sommer, M. H. 2007. "The Adjudication of Wife-selling in Qing County Courts: 272 Cases from Ba, Nanbu, and Baodi Counties." Memo. Stanford University.
- Su, J. and J. He. 2010. "Does Giving Lead to Getting? Evidence from Chinese Private Enterprises." *Journal of Business Ethics*, 93(1): 73 -90.
- Tabellini, G. 2008. "The Scope of Cooperation." *Quarterly Journal of Economics*, 123: 905-950.
- Tabellini, Guido. 2008a. "Institutions and Culture." Presidential Address at the European Economic Association in Budapest, *Journal of the European Economic Association*, .
- Telford, T.A. 1986. "Survey of Social Demographic Data in Chinese Genealogies." *Late Imperial China*, 7(2): 118-146.
- Todd, E. 1990 : *L'invention de l'Europe*. Paris: Seuil
- Thøgersen, S. 2002. "Book Review: Village Economy and Culture in Contemporary China." *Modern China*, 28: 253-274.
- Tsai, L.L., 2007. *Accountability without Democracy*. Cambridge, UK: Cambridge University Press.
- Voigtländer, N. and H.J. Voth. 2009. "The Three Horsemen of Growth: Plague, War and Urbanization in Early Modern Europe." CEPR, DP7275.
- Whyte, M.K. 1995. "The Social Roots of China's Economic Development." *China Quarterly* 144: 999-1019.
- Whyte, M.K. 1996. "The Chinese Family and Economic Development: Development: Obstacle or Engine?" *Economic Development and Cultural Change*, 45(1): 1-30.

- Wolf, C. 1955. "Institutions and Economic Development." *The American Economic Review*, 45(5): 867-883.
- Wood, R. 2003. "Urbanization of Europe and China during the second millennium." *International Journal of Population Geography*, 9: 215-227.
- Zelin, M., J. Ocko and R. Gardella (eds.). 2004. *Contract and Property in Early Modern China*, Stanford, CA: Stanford University Press.
- Zhang, M. 2008. "From Public to Private: The Newly Enacted Chinese Property Law and the Protection of Property Rights in China." *Berkeley Business Law Journal*, 5(2): 317-363.

10 Tables and Figures

Years	to 960	960-1279	1279-1368	1368 - 1644	1644-1911	1911-1949
Period		Song	Yuan	Ming	Qing	Republic
East	0.01	0.02	0.11	0.4	0.24	0.22
South		0.05	0.04	0.3	0.47	0.14
West				0.14	0.73	0.13
North				0.14	0.34	0.52
Source: Telford, 1986, table 10.						
Clan distribution is proxied by (1002) genealogies. We thank Carol H. Shiue for bringing this data to our attention.. North: Hebei, Jilin, Liaoning, Shanxi. East: Anhui, Fujian, Jiangsu, Shandong, Zhejiang. South: Guangxi, Henan, Hubei, Hunan, Shaanxi. West: Guangdong, Guizhou, Jiangxi, Sichuan and Yunnan.						

Table 1: Percentage of regional genealogies tracing the clan's origin to a given period

Year	Cities > 10K		Cities > 40K	
	Europe	China	Europe	China
1000	0	3.1	na	na
1500	5.6	2.8	1.9	na
1600	7.6	4	3.5	1.7
1700	9.2	na	5.2	1.9
1800	10	3.8	5.6	1.7
Sources: Wood, 2003 and Maddison, 2001.				

Table 2: Percent living in cities with above 10K and 40K residents

Table 4 – Trust attitudes: Chinese vs European immigrants

Dependent Variable	(1)	(2)	(3)	(4)
Estimation technique	<i>Trust family</i> OLS	<i>Trust colleagues</i> OLS	<i>Trust neighbors</i> OLS	<i>General Trust</i> Probit
PANEL A				
<i>Chinese vs European</i>	-0.002 (0.022)	-0.055 (0.038)	-0.220*** (0.039)	-0.060*** (0.021)
Observations	17,352	14,677	17,278	17,207
R2	0.015	0.049	0.102	0.026
PANEL B				
<i>Chinese vs European</i>	0.055** (0.023)	0.050 (0.039)	-0.125*** (0.039)	-0.045** (0.021)
<i>Uneasy</i>	-0.073*** (0.008)	-0.137*** (0.012)	-0.122*** (0.011)	-0.019*** (0.006)
Observations	17,325	14,662	17,252	17,182
R2	0.024	0.059	0.109	0.027

Robust standard errors in parenthesis: * significant at 10%; ** significant at 5%; *** significant at 1%

Column (4) reports marginal effects and Pseudo R-squared.

All regressor include the following controls (besides those reported in the Table): Total household income in 2005 Canadian dollar; number of persons in the household; number of children in the household; age; gender; a dummy variable if completed high school, if completed college, if married, if working, if second generation immigrant.

Role of Clan Org	N	Percentage
Dispute resolution	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%
Source: GSS06 (qf13)		

Table 5: What do clan organizations currently do?

Dependent variable	(1)	(2)	(3)	(4)
		<i>Urbanization around 1600</i>		
<i>Extended family</i>	-0.118*** (0.035)	-0.124** (0.040)	-0.120** (0.041)	-0.110** (0.040)
<i>Sea</i>	11.709*** (2.076)	14.476** (5.476)	14.966** (5.459)	14.800** (5.947)
<i>Atlantic Ocean</i>	-3.995* (2.177)	-7.130 (5.878)	-7.420 (5.942)	-7.716 (7.068)
<i>Longitude</i>		-0.407 (0.375)	-0.463 (0.359)	-0.442 (0.361)
<i>Population density</i>			14.545*** (4.095)	8.074 (6.894)
<i>Capital</i>				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 6: Regression results, Europe

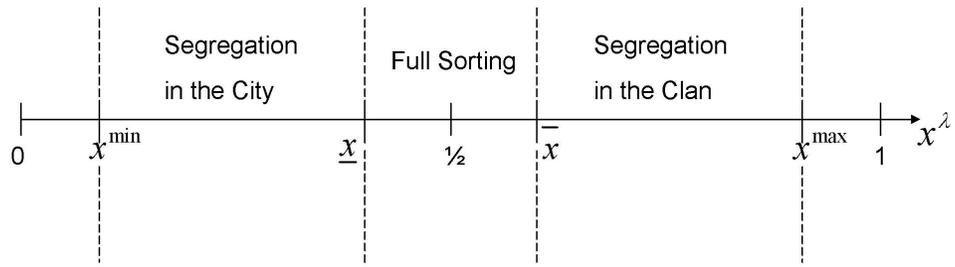


Figure 1: Equilibrium configuration.

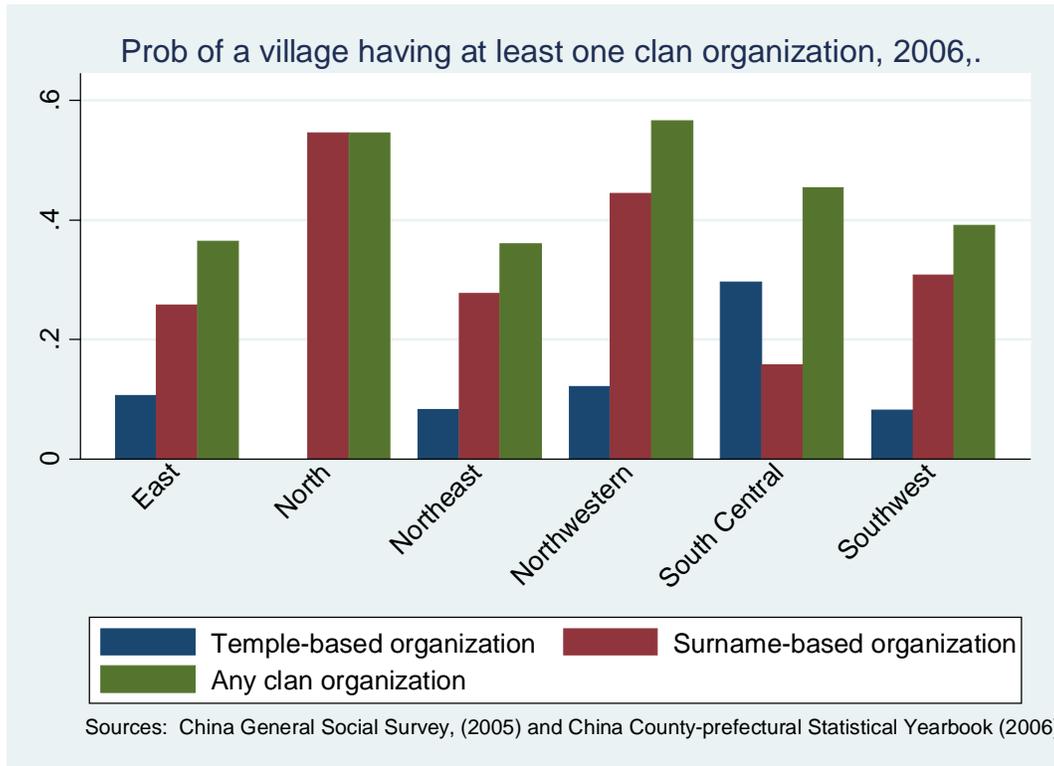


Figure 2: Clan organization by regions.

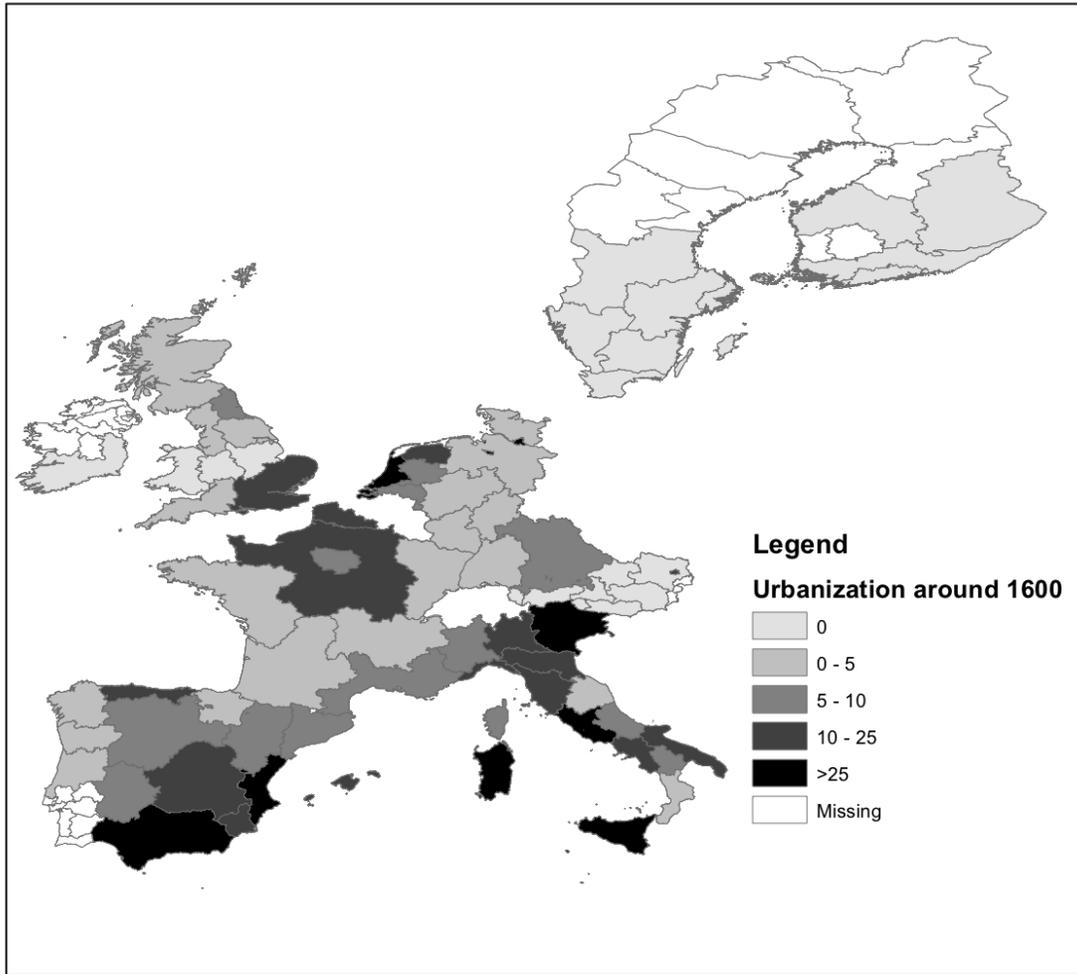


Figure 3a: Urbanization around 1600

Note to figure 2a: 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.

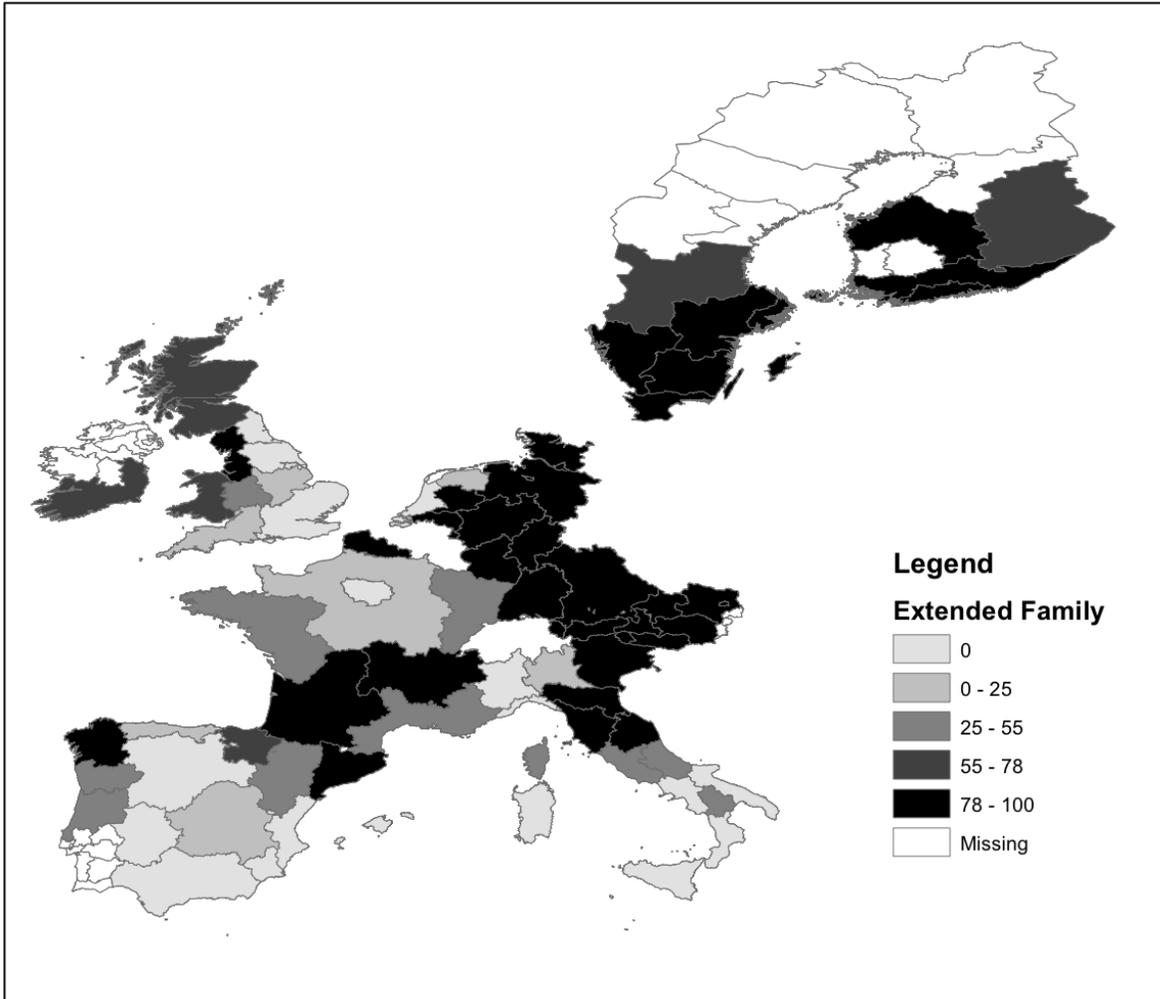


Figure 3b: Extended Family.

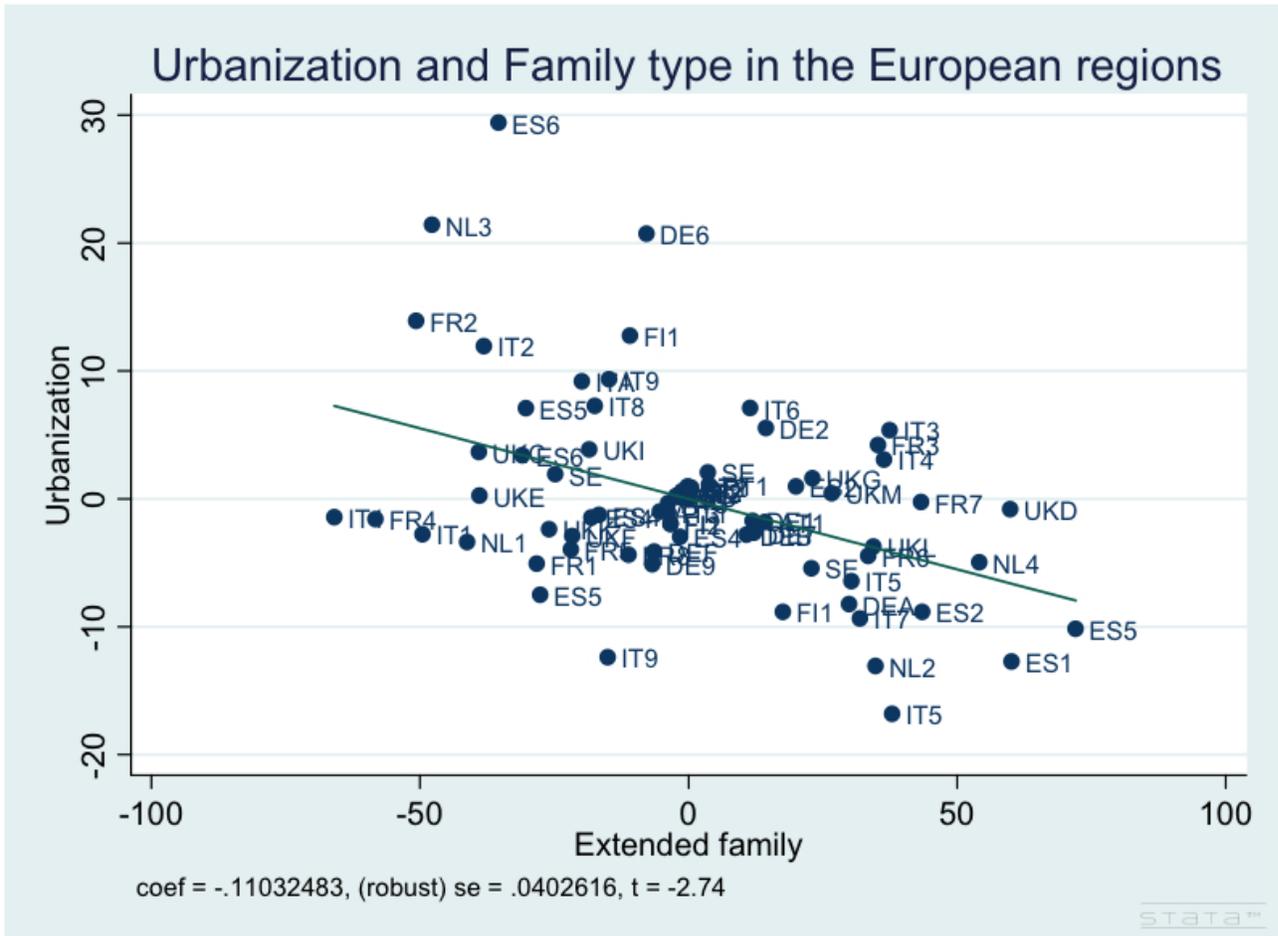


Figure 4: Urbanization and Family type in the European regions.