Journal of Economic Behavior and Organization xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Kinship, fractionalization and corruption☆

Mahsa Akbari[a], Duman Bahrami-Rad[b], Erik O. Kimbrough[c,*]

[a] Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BA V5A1S6, BC Canada
[b] Department of Human Evolutionary Biology MCZ 533C, Harvard University, 11 Divinity Avenue, Cambridge, MA 02318, USA
[c] Smith Institute for Political Economy and Philosophy, Chapman University, One University Drive, Orange, CA 92866, USA

A R T I C L E I N F O

Article history:
Received 23 December 2018
Revised 22 July 2019
Accepted 24 July 2019
Available online xxx

JEL classification:
D7
J1
K4
N3

Keywords:
Corruption
Fractionalization
Institutions
Mating patterns
Consanguinity

A B S T R A C T

We examine the roots of variation in corruption across societies, and we argue that marriage practices and family structure are an important, overlooked determinant of corruption. By shaping patterns of relatedness and interaction, marriage practices influence the relative returns to norms of nepotism/favoritism versus norms of impartial cooperation. In-marriage (e.g. consanguineous marriage) generates fractionalization because it yields relatively closed groups of related individuals and thereby encourages favoritism and corruption. Out-marriage creates a relatively open society with increased interaction between non-relatives and strangers, thereby encouraging impartiality. We report a robust association between in-marriage practices and corruption both across countries and within countries. Instrumental variables estimates exploiting historical variation in preferred marriage practices and in exposure to the Catholic Church's family policies provide evidence that the relationship could be causal.

1. Introduction

Norms of solidarity, fidelity and self-sacrifice in favor of kin, tribe and clan have often been praised as virtues, but these virtues may become vices when they conflict with the abstract rules and formal institutions of a modern political and economic system. In particular, favoring kin *at the expense of others* may lead to corruption, disrupting or subverting impartial institutions and hampering economic development. In this paper, we argue that a history of consanguineous mating practices generates fractionalization between local, sub-ethnic groups, increases incentives for in-group favoritism, and thus encourages corruption.

Previous studies of corruption and its effects on growth have explored the idea that ethnic heterogeneity (and concomitant fractionalization) may cause corruption when individuals favor members of their own ethnic group. Ethnic fractionalization measures are defined as the probability that two randomly drawn individuals from a country's population belong to

☆ Kimbrough would like to thank the SSHRC Insight Grants Program (435-2015-0798) and the Simon Fraser University Teaching and Learning Centre Development Grants Program for funding this research. We would like to thank three anonymous referees, Quamrul Ashraf, Chris Bidner, Greg Dow, Simon Halliday, Alex Karaivanov, Jonathan Schulz, Bart Wilson, and participants at numerous seminars and conference presentations for helpful comments. We also thank Babak Jahanshahi and Erasmo Papagni for helping us collect data on Italy. All remaining errors are our own.

* Corresponding author
E-mail address: ekimbrou@chapman.edu (E.O. Kimbrough).

https://doi.org/10.1016/j.jebo.2019.07.015
0167-2681/

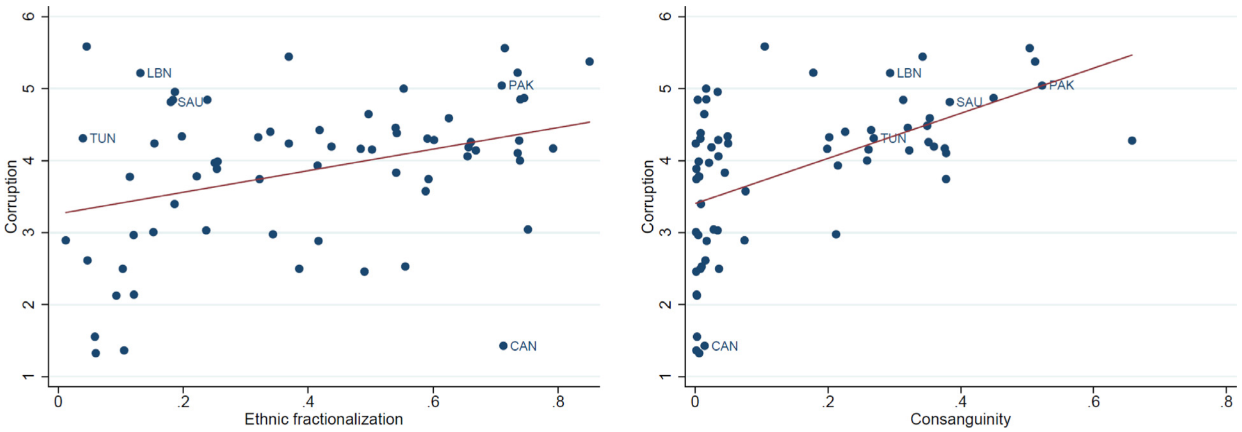

Fig. 1. Corruption (ICRG), ethnic fractionalization (Alesina et al., 2003), and consanguinity (Bittles and Black, 2015). The sample of countries with corruption data is restricted to those with available consanguinity rates. Corruption index (1–6) is reversed (6-1) so that higher values represent higher corruption.

two different groups (Bruk, 1964; Alesina et al., 2003). Mauro's (1995) influential study used ethnic fractionalization as an instrumental variable for corruption. Since then, several studies have investigated whether fractionalization causes corruption, with mixed results. In support, Easterly and Levine (1997) found that ethnic fractionalization is positively correlated with corruption; La Porta et al. (1999), Treisman (2000) and Alesina et al. (2003) also found that fractionalization has a reduced-form relationship with corruption but reported non-robust results when controlling for other variables such as per capita income. However, Serra (2006) and Elbahnasawy and Revier (2012) found no significant effect of fractionalization on corruption. In addition to the cross-country studies, Glaeser and Saks (2006) and Dincer (2008) found a significant relationship between ethnic heterogeneity and corruption across US states. These contradictory results have encouraged skepticism (see e.g. Chuah et al., 2013).

A typical regression model from the cross-country empirical studies is as follows: $C = \alpha + \beta EF + \gamma X + u$, where C is a corruption index, EF is an ethnic fractionalization index, and X is a set of independent variables. The richest specifications in La Porta et al. (1999) and Alesina et al. (2003) include legal origins, religion, latitude, per capita income, country size and regional dummies. We contribute to this literature by highlighting an overlooked aspect of fractionalization in these analyses, what we call sub-ethnic fractionalization.

To see how sub-ethnic fractionalization can help account for some puzzling observations on the relationship between ethnic (and linguistic) fractionalization and corruption, consider Fig. 1. Although Tunisia, Saudi Arabia and Lebanon are relatively homogeneous in terms of ethnic fractionalization ($EF < 0.2$, according to Alesina et al., 2003), they are highly fractionalized due to the presence of and competition between other close-knit kin-based and local groups such as extended families, tribes, clans, and religious groups (see e.g. Lewis, 2014, on clan structures in the Gulf of Aden). We argue that distinctive extended family structures and mating patterns generate sub-ethnic fractionalization and can help account for corruption in many societies, even those that are ethnically homogeneous. This is reflected in high consanguineous marriage rates in these countries. Consanguineous marriages "historically provided one means of creating and maintaining kinship groups–such as clans, lineages, and tribes" (Greif, 2006, p. 309). For this reason, we use consanguineous marriage rates as a proxy for sub-ethnic fractionalization due to the presence of kin-based groups.

A second puzzle for the view that heterogeneity *per se* causes corruption can be seen among countries at the higher end of ethnic fractionalization in Fig. 1. Although, e.g., both Canada and Pakistan are ethno-linguistically heterogeneous, Canada has effective, impartial institutions; while Pakistan is quite corrupt. As above, the countries differ in the importance of sub-ethnic groups such as extended family, tribe and clan to social and political life. Pashtuns, one of the largest ethnic groups in Afghanistan and Pakistan "are said to having [sic] developed the world's largest tribal society,...[with] sub-tribes, clans and sub-clans down to the local lineages and families" (Glatzer, 2002, p. 3). Viewed through the lens of marriage practice-driven sub-ethnic fractionalization, this observation can be understood: Afghanistan and Pakistan have high consanguineous marriage rates; Canada does not.

Important for our research, a large portion of the observed variation in consanguineous marriage rates has deep historical roots. The Catholic Church has restricted cousin marriage since 500CE, at times extending a ban as far as sixth cousins. The Church's marriage policies produced "a vast range of people, often resident in the same locality, that were forbidden to marry" (Goody, 1983, p. 56) and rendered in-marriage virtually impossible. Under strict consanguinity prohibitions, those living in small groups or remote regions had to emigrate to find a legitimate marriage partner (Cavalli-Sforza et al., 2004). These ties facilitated future marriages, exchanges, and other inducements to impersonal cooperation. The Church's family and marriage policies fundamentally altered social organization by slowly dissolving clan, lineage, and tribe boundaries in Europe, and thereby created pressures favoring the development of impartial norms, large-scale cooperation, and eventually modern institutions (Goody, 1983; Herlihy, 1985; Ekelund et al., 1996; Korotayev et al., 2004; Greif, 2006; Mitterauer, 2010).

Please cite this article as: M. Akbari, D. Bahrami-Rad and E.O. Kimbrough, Kinship, fractionalization and corruption, Journal of Economic Behavior and Organization, https://doi.org/10.1016/j.jebo.2019.07.015

Meanwhile, other societies such as in China and the Middle East mostly remained clan-based (Greif and Tabellini, 2015), where people were bound together by strong ties of loyalty, reinforced by marrying within the clan, usually with cousins (Edlund, 2018). While exogamous choice of partners in Europe served as a model for bureaucratic relations by creating links between unrelated individuals, in-marriage in clan-based societies posed an obstacle to the creation of impersonal relationships by cutting “horizontally across the vertical edifice of the state, undermining the system and producing what in conventional administrative terms is called corruption” (Todd and Garrioch, 1985, p.146).

One way of viewing our hypothesis is as follows: corruption based on kinship ties and dense networks of interaction is straightforward to explain, since favoring kin and in-group members comes quite naturally and is reinforced by marriage practices that sharpen divisions between groups and densify those groups in tribal, clan-based societies. The puzzle is to explain the relative absence of corruption and the development of impartial, impersonal institutions in some places and times. We conjecture that changes in marriage patterns (in particular historical reductions in consanguineous marriage and other in-marriage practices) helped create incentives for the creation of impartial, impersonal institutions as individuals sought substitutes for services that were previously provided in kin- and clan-based networks (e.g. mutual insurance). We show that historical and contemporary data on marriage practices are robust correlates of corruption.

Our empirical analysis combines data from population genetics, corruption, and comparative development studies to test the hypothesis that marriage practice-driven sub-ethnic fractionalization is associated with corruption using both cross-country and within-country regression analysis. First, we collect data on national and regional (across-Italian provinces and across European regions) rates of consanguineous (cousin) marriage. We find that consanguinity rates have a substantial and positive association with corruption, both across and within countries, even after controlling for other “deep” determinants of comparative development.

While we cannot allay all endogeneity concerns with our regression analyses, we report additional results based on historical cultural measures and instrumental variables estimates to provide some evidence that the relationship we identify could be causal. In our cross-country analysis, to complement our analyses using cousin marriage rates, we use data on cousin marriage practices of pre-industrial societies from the Ethnographic Atlas, and following the methods developed by Alesina et al. (2013) for aggregating data about the cultural practices of ethnic groups into country-level statistics, we generate a country-level measure of historical lineage endogamy—defined as a preference for marriage between the children of two brothers (aka parallel cousin marriage). We then regress present day corruption on historical lineage endogamy.

Lineage endogamy, via parallel cousin marriage, unites two people born *within* a single male lineage group, reinforcing and deepening the divisions *between* lineage groups and increasing fractionalization. As marriage becomes increasingly exogamous (e.g. across lineages, villages, regions, etc), it begins to break down between-group barriers and reduce fractionalization. By way of contrast, cross-cousin marriage (i.e. the marriage between the children of a brother and a sister) can be seen as a weak form of exogamy, since such a marriage unites two distinct male lineages. Endogamy (and consanguinity) amplify the incentives for norms of in-group favoritism, while exogamy favors the development of norms facilitating cooperation with outsiders. Thus we expect a connection between lineage endogamy and corruption. This measure also has the advantage of ruling out the possibility of reverse causality, since it pre-dates contemporary measures of development and institutional quality such as corruption. The analysis shows that lineage endogamy is associated with significantly higher levels of corruption after controlling for deep determinants, income per capita, and even the population share of different religious denominations.

Next, relying on arguments regarding the historical influence of the Catholic Church in reducing past (and thereby contemporary) cousin marriage rates and inspired by Schulz (2017), we use data coded by Korotayev (2003) to capture 'deep Christianization' of the Ethnographic Atlas societies. We show that pre-industrial deep Christianization—a 500-year exposure to Christianity—is associated with lower corruption today in both reduced-form and instrumental variable estimations. This provides some support for the hypothesis that the Church's family and marriage policies played a role in weakening kin-based structures, facilitating the growth of impartial institutions. Of course, the evidence is not conclusive since in the absence of a natural experiment in history, it's difficult to establish decisive causal links between the Church's marriage and family policies, marriage practices and corruption.

Since cross-country analyses are fraught with concerns about potential confounding variation, we provide additional support in a within-country analysis using province-level data from Italy. We show an association between cousin marriage rates (measured during the 1950-60s) and today's corruption rates that is robust to inclusion of many variables to which regional differences in Italy are attributed in the literature. We also construct an instrument that captures province-level variation in exposure to the Catholic Church's consanguinity bans by using data on the location and dates of establishment and suppression of Catholic dioceses at half-centuries between 500 and 1500 CE. Exposure to the Catholic Church significantly explains today's corruption rates across Italian provinces in both reduced-form and instrumental variable estimations.

Italy is a uniquely interesting case in that by the 20th century (when we observe consanguinity rates) it was a nearly 100% Catholic country, but some southern parts of Italy, where cousin marriage is most common in our data, were also part of the Islamic Caliphate or Byzantium and the Langobardic kingdoms for centuries. Therefore, both historical measures of exposure to the Catholic Church and contemporary consanguinity rates partly reflect historical military and cultural conquests which carried with them different attitudes towards family and marriage patterns, and in our view, the impact of this variation in marriage practices on patterns of relatedness has helped shape the extent of corruption and the quality of institutions today.

Finally, we also show that the association between corruption and consanguinity remains robust in a sample of European regions (across France, Italy and Spain), all of which are historically Catholic. Overall, our findings suggest that marriage practice-driven sub-ethnic fractionalization is an alternative but important channel through which history can explain variation in present day institutional quality.

Our work is also related to a growing literature on the political and economic consequences of historical variation in social and familial organization (e.g. Woodley and Bell, 2013; Schulz, 2017; Moscona et al., 2017a; 2017b; Enke, 2017). Moscona et al. (2017a,b) provide evidence that other aspects of kinship and social organization help account for levels of inter-group conflict in Sub-Saharan Africa; in particular, they show that groups organized around segmentary lineages are prone to localized conflicts. The most closely related papers by Woodley and Bell (2013) and Schulz (2017) document a negative relationship between rates of consanguineous marriage and levels of democracy. We view this work as complementary to our own, as there is a documented association between corruption and democracy. Moreover, we would suggest that the mechanism we propose, which connects kinship and marriage practices to norms of favoritism, and hence to corruption, may partly underpin the relationship they observe, as kin-based corruption is inimical to democratic institutions. Enke (2017) develops a measure of historical "kinship tightness" along two dimensions: family structure (domestic organization, post-wedding residence) and descent systems (lineages, segmented communities). He reports a negative correlation between his measure of kinship tightness and cooperation with and trust of outsiders using a variety of datasets. Schulz et al. (2018) also develop an index of "kinship intensity" by combining different dimensions of kin-based organization (cousin marriage, polygyny, co-residence of extended families, unilineal descent, and community organization) and show that kinship intensity explains psychological variation across populations. Schulz (2017) and Schulz et al. (2018) also use different measures based on exposure of countries and regions to Catholicism and languages differentiating terms for cousins.[1]

2. Consanguineous marriage, sub-ethnic fractionalization, and corruption

Our study lies in the literature that distinguishes between generalized morality and limited morality (or amoral familism, Banfield, 1958) and their consequences. Limited morality is the extreme reliance on a narrow circle of family, friends or relatives; outside this circle, harming and cheating are allowed and frequent. In this narrow circle, people are raised to trust in-group members only. They are also taught to distrust people outside the circle, which hampers cooperation and exchange with strangers and outsiders, and as a result, impedes the development of formal institutions. Generalized morality is characterized by respect for abstract individuals and their rights, generalized trust and loyalty to general rules, which facilitates large-scale cooperation. However, the underlying mechanism that determines whether a society adopts limited morality has been attributed to collectivism versus individualism (Yamagishi and Yamagishi, 1994; Yamagishi et al., 1998), strong versus weak family ties (Ermisch and Gambetta, 2010; Alesina and Giuliano, 2011; 2014), and clan versus corporation (Greif and Tabellini, 2015; Greif, 2006). We contribute to this literature by providing evidence for the latter.

Clan-based societies differ from corporation-based societies in their basic units of cooperation—the clan and the corporation—which are two distinct social organizations and imply different notions of morality. A corporation is a voluntary association between unrelated individuals established to pursue common interests. Modern institutions, the modern state, and the rule of law evolved gradually in Europe from pre-existing corporations such as communes, guilds, universities and business associations—which were "built on impersonal rules consistent with generalized morality, and more appropriate to enforce cooperation between unrelated individuals" (Greif and Tabellini, 2015, p.3). This tradition in which individuals unrelated by blood cooperated based on generalized morality facilitated their integration into the state administration and bureaucracy (Greif and Tabellini, 2015). By contrast, the clan consists of families united by common descent who share reciprocal loyalty and obligations to each other which are reinforced by the common practice of marrying within the clan, usually cousins (Edlund, 2018). By increasing the strength of ties between clan members, in-marriage raises the relative returns to limited morality and thereby encourages corruption in the context of impersonal institutions, such as state bureaucracies.

Cousin marriage also increases biological relatedness at a local level, thereby directly altering the incentives for corruption and favoritism (Hamilton, 1964; 1975). For instance, the offspring of a consanguineous marriage (e.g. marriage between two first cousins) share more genes with their parents, siblings and cousins than the offspring of a marriage between two unrelated individuals. From the point of view of biological theory, all else equal, the former family has stronger incentives for favoritism and corruption at the expense of non-relatives than the latter.

In addition, cousin marriage can directly influence the evolution of social norms. Many groups extend altruistic norms to affines, friends and other less-related individuals by adopting cultural practices that create fictive kinship, and it has been argued that these practices piggy-back on, or hijack, the evolved mechanisms for kin cooperation and extend them to non-kin (Henrich, 2015). In particular, different marriage patterns also typically result in different patterns of social interaction, reflecting complementarities between marriage practices and other social structures. Out-marriage encourages (indeed, requires) interaction with non-kin and strangers. By contrast, in-marriage is associated with extensive interaction among local (and more closely related) in-group members (e.g. kin, clan, tribe), and relatively less interaction with strangers. Seen in this light, in-marriage practices generate another kind of *fractionalization* at a level of granularity finer than the ethnic or

[1] The primary analyses that make up this article were undertaken prior to the publication of the above-mentioned papers, and at that time, we were aware only of the ongoing related work by Schulz (2017) and the earlier paper by Woodley and Bell (2013).

Table 1
Religious attitudes to consanguineous marriage (Bittles, 2012, p. 14).

Religion	Sect	Attitude toward Cousin Marriage
Judaism	Ashkenzi	Permitted
	Sephardi	Permitted
Christianity	Coptic Orthodox	Permitted
	Eastern Orthodox	Proscribed
	Protestant	Permitted
	Roman Catholic	Diocesan approval req.
Islam	Sunni	Permitted
	Shia	Permitted
Hinduism	Indo-European	Proscribed
	Dravidian	Permitted
Buddhism		Permitted
Sikhism		Proscribed
Confucian/Taoist		Partially permitted
Zoroastrian/Parsi		Permitted

linguistic group. We refer to this as *sub-ethnic fractionalization*, and we argue that patterns of increased local relatedness and concomitant intense local interaction diminished the impetus to develop norms of impartial cooperation, instead favoring the development of norms of local favoritism (which may manifest as corruption).

2.1.1. On the historical origins of differences in kin-based institutions

In fact, a famous theory developed by authors like Goody (1983), Herlihy (1985), Ekelund et al. (1996), Greif (2006), Mitterauer (2010) and Greif and Tabellini (2015) traces the divergence between corporation-based societies (such as in Western Europe) and clan-based societies (such as in China and the Middle East) and the differences in moralities back to religious beliefs concerning matters of family and marriage originating in the distant past. In Europe, Christianity encouraged moral obligations toward non-kin by putting emphasis on the individual rather than the family. This encouraged the later prohibitions on the subjects of family and marriage that had an important role in the decline of large kinship groups. The Church discouraged practices that enlarged the family, such as adoption, polygamy, concubinage, easy divorce, remarriage, and most importantly, it also prohibited consanguineous marriages. As the Church extended prohibitions to remoter degrees of consanguineous marriage, clans, lineages, and tribes gradually disappeared throughout Europe such that there is a large and significant negative correlation between the spread of Christianity and the absence of clans and lineage groups in Europe (Korotayev, 2003). With the disappearance of kin-based groups, the European nuclear family structure and individualistic culture emerged in the late medieval period and supported the development of corporations (Goody, 1983; Greif, 2006).

By contrast, societies such as China and those in the Middle East have been clan-based for most of their history. In China, Confucianism and post-Song Chinese Buddhism emphasized moral obligations towards kin. Confucius (551-479 BCE) argued that one has the moral obligation to protect kin, even one who broke the law and harmed others. Confucius viewed marriage as a “union of two surnames, in friendship and in love” (Greif and Tabellini, 2015, p.20,28). Kin-based groups such as extended family and tribes and a persistent preference for cousin marriage (at rates as high as 50% of marriages) mark Middle Eastern societies even today (Greif, 2006; Todd and Garrioch, 1985; Bittles, 2012). Here again, religion is thought to have an important role. Islam served as a vehicle for the spread of (parallel) cousin marriage, a custom of Arab populations (Courbage and Todd, 2014).

The wide diversity in attitudes towards consanguinity in human societies partially originates in religious beliefs due to the common jurisdiction of religious institutions over marriage. Table 1 shows some of the diversity of religious attitudes toward cousin marriage around the world, and Appendix A summarizes the history of Christian and Islamic attitudes toward consanguinity.

In fact, marriage norms may persist, even as religious attitudes change. As one example, consanguineous marriage has long been prevalent in parts of Italy, despite it being an almost entirely Catholic country. Cavalli-Sforza et al. (2004) suggest this may be a result of persistent cultural norms imported during the Arab conquest of southern Italy over 1000 years ago. Going the other direction, majority-Protestant countries mostly legalized cousin marriage after centuries of living under the Catholic ban. Nevertheless, consanguinity remains rare in those countries. In the United States, cousin marriage is illegal in 25 states, though its frequency remains low even where it is legal.

Of course, not all consanguineous marriage is driven by religious preference. For instance, consanguineous mating can be caused by *population division* due to geography. As populations migrated around the world historically, they became isolated from one another due to vast distances and geographic barriers such as mountains, deserts and oceans that were only recently broken down by transportation technologies. Due to isolation, it was virtually impossible not to marry a relative.

Others have argued that consanguinity is a cultural adaptation to social and ecological circumstances. In his history of the family, Goody (1983) suggests that consanguinity may be a property and wealth-preserving response to gender-egalitarian inheritance rules, which would result in the diffusion of property under out-marriage; see Bahrami-Rad (2018) for

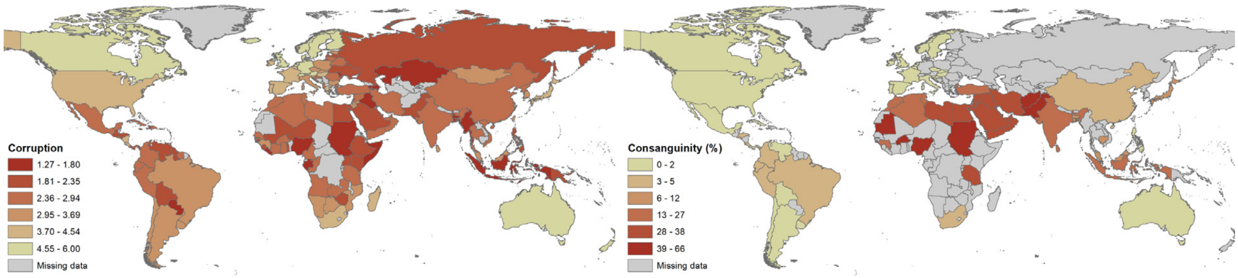

Fig. 2. Corruption (ICRG) and consanguinity (Bittles and Black, 2015) around the world. Grey colored areas indicate missing data.

supporting evidence. A few studies have examined the causes and consequences of consanguinity in small-scale societies using detailed genealogies to directly measure relatedness. Walker and Bailey (2014) show that among forager peoples, such marriages are rare due to norms of exogamy and fission/fusion dynamics that disperse kin across groups, but among agropastoralists, particularly those that practice polygyny, the practice is more common, with average spousal relatedness rising as high as $r = 0.18$ (almost 50% greater than first cousins, $r = 0.125$). Using $\log(\text{surviving children})$ as a measure of fitness, estimates in Bailey et al. (2014) suggest that these marriage practices may be adaptive, with fitness maximized for moderate consanguinity among agropastoralists and with minimal consanguinity among foragers. Other evidence from Hoben et al. (2010) suggests that consanguinity may be more prevalent near the equator since it can raise the frequency of homozygosity for adaptive recessive mutations that defend against diseases and parasites, which are also more prevalent in warmer climes.

Regardless of its origins, wherever it is practiced, consanguineous marriage directly increases local relatedness and encourages sub-ethnic fractionalization, thereby altering the returns to norms of favoritism and norms of impartial cooperation. Thus, variation in consanguinity rates facilitates a test of our main hypothesis: that sub-ethnic fractionalization causes corruption.

3. Empirical strategy and results

To provide evidence for our hypothesis, we present data from cross-country regressions and within-country regressions. Each of our analyses has limitations, and none of them can wholly address identification or endogeneity concerns, but our goal is that by approaching the problem at various levels of granularity we can provide a set of robust, complementary tests of the main hypothesis: that sub-ethnic fractionalization causes corruption.

In our cross-country and within-country analyses, we employ data on consanguineous marriage as our proxy for sub-ethnic fractionalization, and we regress measures of corruption on consanguinity. As discussed below and detailed in Appendix B, Bittles and Black (2015) have collected data on rates of consanguineous marriage (among 2nd cousins or closer relatives) at the country level from a multitude of sources, including surveys, public health studies, and church records. While our cross-country consanguinity sample is neither random nor representative, it covers the large majority of the global population. For our within-country analysis, the data on consanguinity has more complete coverage, as Cavalli-Sforza et al. (2004) have collected data on consanguinity in Italian provinces. All data are described in detail in Appendix B.

3.1. Cross-country analysis

Fig. 2 displays average International Country Risk Guide (ICRG) corruption data from 1984 to 2011 alongside consanguinity data from Bittles and Black (2015). We follow La Porta et al. (1999) and Alesina et al. (2003) and in our main analyses use ICRG corruption scores that have been widely used in corruption studies (see also Alesina and La Ferrara, 2002; Fisman and Gatti, 2002; Persson et al., 2003; Brunetti and Weder, 2003). However, we also check the robustness of our regression results to alternative measures of corruption in the Appendix. The corruption data is collected by the ICRG, which produces annual ratings of corruption levels via surveys of country experts. The indicator ranges from 0 to 6 with higher values implying lower corruption. The indicator accounts for financial corruption in the form of demands for special payments and bribes connected with import and export licenses, exchange controls, tax assessments, police protection, or loans, and also actual or potential corruption in the form of excessive patronage, nepotism, job reservations, favor-for-favors, secret party funding, and suspiciously close ties between politics and business (Howell, 2011).

Although we have consanguinity data for 72 countries, in our primary analyses, there are 65 countries for which we have the full set of covariates used in our main regression analyses. In this sample, we find a negative and highly significant correlation between corruption and consanguinity (Pearson's $r = -0.51$, p-value < 0.001, N = 65).

While the raw correlation is strong, we also conducted a series of OLS regressions controlling for relevant confounds and alternative explanations.[2] Our empirical strategy follows La Porta et al. (1999) and Alesina et al. (2003) who attempted to

[2] Results are robust to using Tobit to account for the dependent variable being restricted to the interval [0,6].

address the endogeneity of corruption by focusing on "(reasonably) exogenous sources of variation" (La Porta et al., 1999, p. 223) in the *economic, geographical, political and cultural* characteristics of countries. Hence, we do not include contemporary political variables or variables capturing public policy. Instead, our analysis focuses on "more fundamental, or at least historically predetermined" variables (La Porta et al., 1999, p. 230; see also Treisman 2000, p. 409).

The most obvious economic heterogeneity across countries that can affect corruption is economic development, but development is almost certainly endogenous to corruption. While there is evidence that poor countries are perceived to be more corrupt than rich ones, so that per capita income is a potential determinant of corruption, corruption itself can reduce per capita income (see e.g. Mauro, 1998; Tanzi and Davoodi, 1998; Campos et al., 1999; Lambsdorff, 2003). As noted by Treisman (2000), one exogenous variable that is correlated with economic development but is unaffected by corruption is a country's latitudinal distance from the equator (indeed, log GNI per capita averaged over 1984–2011 correlates with latitude, Pearson's $r = 0.65$, p-value < 0.001, N = 65).[3] In addition to latitude, following Alesina et al. (2003), we also include regional dummies for Sub-Saharan Africa, East Asia and Pacific, Latin America and Caribbean and country size (population). Following the literature, we also include ethnic fractionalization and legal origins as sources of exogenous variation in country-level political characteristics as well as geographic variables from Ashraf and Galor (2013) including ruggedness, soil suitability for agriculture, mean elevation, mean temperature, mean precipitation, percentage of population living in tropical and subtropical climate, percentage of population living in temperate climate, and percentage of land near waterways. We report heteroskedasticity robust standard errors.

3.1.1. Main findings

Table 2 displays our first set of regressions, with the ICRG corruption index as the dependent variable. A full description of the variables is presented in Appendix B, Table B.1, and the omitted legal origin dummy in the regressions is the British one. Column (1) presents the simplest model without controlling for latitude or income per capita. Column (2) presents our basic regression model described above and without consanguinity, inspired by La Porta et al. (1999) and Alesina et al. (2003), which includes a set of historically predetermined and exogenous economic, geographical and political variables. In column (3), we run the same regression using only the sample of countries for which consanguinity data exists and find qualitatively similar results. The main results from these regressions are consistent with previous findings: compared to those in common law countries, governments in countries of socialist legal origin are more interventionist, and thus the socialist legal origin is associated with higher corruption. Scandinavian legal origin, on the other hand, is associated with lower corruption.

In column (4), our baseline model, we include consanguinity rates to account for sub-ethnic fractionalization, and the estimated coefficient is significant at the 1% level. The estimate implies that a 1 standard deviation increase in consanguinity is associated with an increase in corruption by 0.6 standard deviations. As the authors along with Treisman (2000) note, ethnic fractionalization has a reduced form relationship with corruption and is not typically significant after controlling for per capita income and latitude, while consanguinity remains significant even including both income per capita and latitude in column (6).

While Alesina et al. (2003) note the difficulty of disentangling the independent effect of ethnic fractionalization from income per capita and latitude because all are highly correlated, their ethnic fractionalization index is insignificant in all specifications, even in the simplest regression excluding controls for income per capita and latitude, as in column (1).

Religion. Religion is potentially relevant to our analysis in two ways: through an indirect effect on corruption via consanguinity and through a direct effect on corruption as discussed in previous work (see La Porta et al., 1999; Alesina et al., 2003; Treisman, 2000).

A preference for consanguineous marriage is common in the Islamic world, and thus consanguinity rates and the percent of the country practicing Islam are highly correlated (Pearson's $r = 0.78$, p-value < 0.001, N=65).[4] In contrast, Catholicism has imposed a long-standing ban on consanguineous marriage, and this is evinced by a strong negative correlation between the share of a country practicing Catholicism and consanguinity (Pearson's $r = -0.65$, p-value < 0.001, N=65). Finally, while Protestant religions do not officially ban consanguineous marriage, the frequency is quite low, and we find a large negative correlation between a country's share of Protestants and consanguinity (Pearson's $r = -0.27$, p-value $= 0.029$, N = 65).

La Porta et al. (1999) and Treisman (2000) offered reasons why Protestantism may be a cultural deterrent to corruption, beyond its relationship to consanguinity. They argue that Protestantism is associated with attitudes such as Weber (1958)'s "Protestant work ethic" and a separation of church and state, both of which may have been conducive to growth and quality government, while countries that were predominantly Catholic or Muslim during this period were relatively more insular, hierarchical and interventionist (see La Porta et al., 1999, p. 229).

In column (5) of Table 2, we include variables indicating the share of each country's population that practices Protestantism, Catholicism, and Islam; the excluded category is "other religions". We confirm La Porta et al.'s (1999) and

[3] One possible indirect connection is through an effect of latitudinal distance from the equator on consanguinity rates. As noted above, Hoben et al. (2010) argue that relative parasite prevalence near the equator may raise the returns to consanguinity by raising the frequency of homozygosity for adaptive recessive mutations (e.g. parasite immunities). We find a high correlation between distance from the equator and consanguinity which provides further reason to control for latitude in our regressions (Pearson's $r = -0.30$, p-value $= 0.015$, N = 65).

[4] Note, however, that we find a significant correlation between consanguinity and corruption even if we focus only on minority-Muslim countries (Pearson's $r = -0.37$, p-value $= 0.012$, N=46).

Table 2
Regression analysis of the relationship between consanguinity and corruption. Higher values of the dependent variable imply lower corruption. Full specification reported in Table C.1 of Appendix C.

VARIABLES	(1)	(2)	(3) Restricted sample	(4) Basic model	(5) and religion	(6) both Income and Latitude	(7) and religion
Consanguinity				−3.874***	−2.438**	−2.852**	−1.743
				(0.999)	(1.143)	(1.242)	(1.416)
				[0.785]	[0.873]	[1.142]	[1.155]
Ethnic fractionalization	0.232	0.271	−0.380	−0.140	−0.171	−0.352	−0.339
	(0.480)	(0.505)	(0.614)	(0.447)	(0.416)	(0.489)	(0.477)
Latitude		0.628	−3.566	−4.349**	−1.711	−3.782*	−1.307
		(1.747)	(2.357)	(2.092)	(2.114)	(2.081)	(2.129)
Log GNI per capita						0.466	0.395
						(0.344)	(0.343)
Protestant					0.848		0.645
					(0.946)		(1.024)
Catholic					−0.743**		−0.836**
					(0.349)		(0.359)
Muslim					−1.547***		−1.533***
					(0.437)		(0.434)
Log population	−0.041	−0.039	−0.098	−0.126	−0.301	−0.098	−0.273
	(0.113)	(0.114)	(0.204)	(0.182)	(0.204)	(0.166)	(0.186)
Africa	−0.196	−0.172	0.052	0.446	−0.104	0.539	0.022
	(0.251)	(0.262)	(0.538)	(0.365)	(0.357)	(0.387)	(0.371)
East Asia	−0.082	−0.069	−0.274	−0.643	−0.773*	−0.570	−0.729*
	(0.341)	(0.335)	(0.489)	(0.433)	(0.415)	(0.402)	(0.377)
Latin America	−0.230	−0.207	−0.185	−0.895**	−1.139***	−0.899**	−1.095***
	(0.261)	(0.265)	(0.430)	(0.369)	(0.327)	(0.365)	(0.316)
Socialist legal origin	−1.362***	−1.373***	−0.895**	−0.989***	−0.907***	−0.846**	−0.794***
	(0.285)	(0.287)	(0.360)	(0.318)	(0.285)	(0.323)	(0.287)
French legal origin	−0.080	−0.083	−0.106	−0.047	0.293	0.037	0.362
	(0.174)	(0.174)	(0.266)	(0.227)	(0.227)	(0.252)	(0.255)
German legal origin	0.442	0.440	−0.019	0.122	−0.072	−0.048	−0.288
	(0.427)	(0.420)	(0.415)	(0.388)	(0.470)	(0.417)	(0.506)
Scandinavian legal origin	1.155***	1.110***	0.825*	0.916*	−0.490	0.822*	−0.484
	(0.368)	(0.369)	(0.478)	(0.458)	(0.846)	(0.452)	(0.841)
	(1.002)	(1.749)	(2.439)	(1.833)	(1.606)	(2.355)	(2.121)
Geographical variables	yes	yes	yes	yes	yes	yes	yes
Observations	126	126	65	65	65	65	65
R-squared	0.653	0.654	0.680	0.770	0.827	0.781	0.834

Robust standard errors in parentheses. Conley standard errors in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Treisman's (2000) finding that the proportion of Protestants in a country's population is associated with lower corruption relative to other religious groups. Moreover, Islam and Catholicism are both associated with higher corruption.

Although the coefficient of consanguinity is smaller, it remains highly significant. However, when both latitude and income per capita are controlled for in column (7), consanguinity becomes insignificant. Note that this regression on 65 observations includes 24 covariates. Moreover, income per capita is endogenous to corruption. However, in subsequent sections, using a different measure of consanguinity that increases our sample size, we see that consanguinity is again significant in regressions analogous to that of column (7), even when controlling for religion.

Alesina et al. (2003) replication. As a robustness check, we replicate two specifications from Alesina et al. (2003). They report a significant correlation between their ethnic fractionalization index and corruption, even after controlling for legal origins, which we replicate in regression (1) of Table 3. In column (2), we replicate their analysis including only those countries for which we have consanguinity data. Ethnic fractionalization remains significant at the 10% level. Including consanguinity in specifications (3) and (4) causes the effect of fractionalization to disappear while consanguinity is significant. Following a second analysis from Alesina et al. (2003), we also introduce additional control variables in Table C.2 of Appendix C. Consanguinity is still significant.[5]

Additional robustness checks. The results are robust to using Conley (1999) standard errors to account for spatial dependence with different cutoffs. For example, in Table 2, the Conley standard errors of consanguinity with a cutoff of 30 decimal degrees is reported in the brackets. The results also hold when we estimate a spatial autoregressive model (SAR).[6]

[5] Our replications of previous studies give slightly different results for Alesina et al. (2003) in regression (1) of Table 3, and Table C.2 of the appendix, possibly because we used different data sources for population and regional dummies.

[6] For example, allowing for spatial lags of the dependent variable, independent variable, and error in the basic model of Table 2, SAR estimates yield a significant 'direct effect' of consanguinity within a country, with a comparable magnitude to our baseline estimates: -3.207 (p-value=0.000). However,

Please cite this article as: M. Akbari, D. Bahrami-Rad and E.O. Kimbrough, Kinship, fractionalization and corruption, Journal of Economic Behavior and Organization, https://doi.org/10.1016/j.jebo.2019.07.015

Table 3
Model specification Table (13), column (2) from Alesina et al. (2003), including consanguinity.

VARIABLES	(1) Alesina (2003)	(2) Alesina (2003) restricted sample	(3) with Consanguinity	(4) with Consanguinity without EF(2003)
Ethnic fractionalization	−2.498**	−2.777*	−0.183	
	(0.981)	(1.394)	(1.196)	
Consanguinity			−10.382***	−10.493***
			(1.579)	(1.274)
Log population (1960)	0.076	0.203	−0.259	−0.253
	(0.260)	(0.340)	(0.244)	(0.249)
Africa	−0.902	0.202	1.475**	1.423**
	(0.569)	(1.061)	(0.709)	(0.598)
East Asia	−1.650**	−1.484	−2.249*	−2.280*
	(0.661)	(1.355)	(1.244)	(1.220)
Latin America	−2.127***	−1.101**	−2.888***	−2.921***
	(0.496)	(0.523)	(0.578)	(0.523)
Constant	7.295***	6.441***	9.112***	9.045***
	(1.052)	(1.485)	(0.996)	(0.995)
Observations	122	64	64	64
R-squared	0.278	0.172	0.519	0.519

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Since the data on consanguinity were collated from 448 studies over many decades, we report additional analysis in Table C.3 in Appendix C that addresses data collection dates. We include the variable of average data collection date of countries weighted by sample size to the baseline regression. Moreover, we run the baseline regression with different consanguinity rates obtained from four groups of studies (with data sampled after 1922, after 1950, after 1960, and after 1970). The results all include a highly significant coefficient on consanguinity.

To make sure our results are robust to alternative measures of corruption, in Table C.4 in Appendix C, we estimate our baseline specification using data on three alternatives: 1) the corruption perception index (1998–2015) from Transparency International, 2) a measure of control of corruption (1996–2017) from the World Bank Group (Kaufmann et al., 2010), and 3) Transparency International's historical data[7] (1980–1985). In all cases, consanguinity remains statistically (at the 5% level) and economically significant.

Although we cannot address all endogeneity concerns with our regression analyses, in Table C.5 in Appendix C, we also check robustness of our baseline regression to inclusion of some other potentially confounding deep-rooted factors highlighted in the literature: family ties (Ermisch and Gambetta, 2010; Alesina and Giuliano, 2011; 2014), trust (Francois and Zabojnik, 2005; Uslaner, 2005; Tabellini, 2010), and genetic diversity (Ashraf and Galor, 2013).[8]

Historical lineage endogamy. One marriage practice encouraging fractionalization via tight kin networks is lineage endogamy, which is a common but relatively extreme form of in-marriage. It is usually organized through the practice of parallel cousin marriage. A parallel cousin marriage consists of marriage between the children of two brothers, who are both members of their father's lineage. This has important implications for social organization, especially in unilineal descent systems where people trace their ancestry back through one side of the family, for example, through the male line in patrilineal societies. Lineage endogamy increases fractionalization, creating overlapping interests between members of the same lineage, and thereby intensifies norms favoring cooperation among them. By comparison, lineage exogamy (marriage to outsiders or even to cross-cousins) creates relatively more competing interests between the two sides of the family and encourages the development of norms facilitating cooperation across lineages (i.e. with outsiders). Moscona et al. (2017b) identify unilineal organization as an important source of conflict between groups in Sub-Saharan Africa.

Lineage endogamy cannot be distinguished in our data on consanguinity rates, but historical data helps us to identify ethnic groups which traditionally practiced it. One source of such historical data used in the economics literature is the Ethnographic Atlas[9] data on 1291 pre-industrial societies. One advantage of this data for our purpose is that it is sampled from pre-industrial societies, and it therefore predates any measures of corruption, which are defined based on modern bureaucratic concepts. To aggregate information about the historical data from the Ethnographic Atlas ethnicities into country-level measures, we follow the method introduced by Alesina et al. (2013). See Alesina et al. (2013) for more details of the

the indirect or 'spillover effect' to other countries (proportional to the inverse of distance between countries) is insignificant: -2.325 (p-value=0.169). Full estimates available upon request.

[7] Available at http://www.icgg.org/corruption.cpi_olderindices.html.

[8] Another potential confound is individualism. However, the most widely used measure of individualism comes from Hofstede (2001), and this measure is based in part on responses to a question about whether it is moral for a boss to offer a job to a relative. This obviously overlaps with our dependent variable ($r = 0.7$). Therefore, we don't report the regression result including individualism where our coefficient of interest is insignificant.

[9] Murdock (1962–1971); Korotayev et al. (2004); Bondarenko et al. (2005); Kirby et al. (2016).

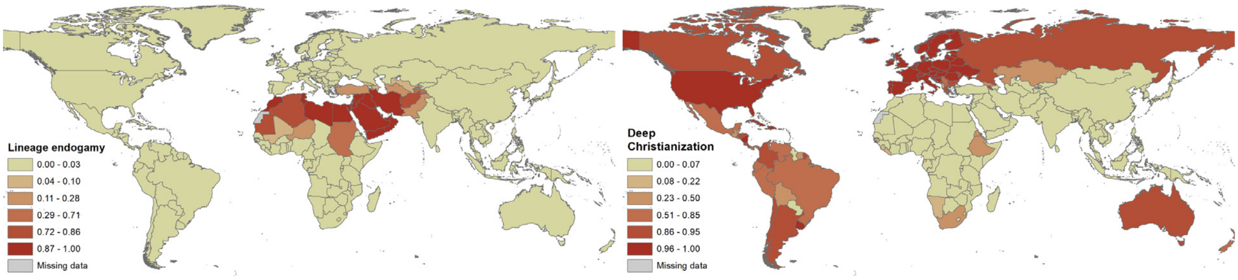

Fig. 3. (Left) Lineage endogamy: preferential marriages between members of the same lineage, i.e., parallel cousin marriage. (Right) Deep—a 500-years—Christianization. See Fig. C.2 of Appendix C for the maps of the measures in ethnicity-level (before aggregation to the country-level).

methodology used in the construction of the country-level data.[10] The country level measure for a variable from the Ethnographic Atlas is the country-level population weighted average of the measure for all ethnicities within the country. As a sanity check, we first successfully replicated their measure of historical plow use (see Fig. C.1 of Appendix C). Then, we created our own measure of lineage endogamy.[11] Fig. 3 (left) shows the global distribution of lineage endogamy measured by preference for parallel cousin marriage defined based on entry 26 of the Ethnographic Atlas. Lineage endogamy is highly correlated with the share of Muslims in the population (Pearson's $r = 0.70$, p-value < 0.001, N = 126).

Table 4 reports regression results analogous to Table 2 except that contrary to the regressions with consanguinity rates, here there is no data restriction due to the explanatory variable. Our sample size is instead restricted by the availability of the other covariates. The results indicate that a tradition of lineage endogamy is associated with a roughly 1 unit increase in the corruption index (or 90% of a standard deviation) in the basic model. The coefficient remains significant even after controlling for income per capita in column (2), and religion in column (3). This implies that the variation remaining in consanguinity rates after controlling for religion (most importantly Islam) can still explain the variation in corruption across countries. However, most countries with high rates of lineage endogamy are in North Africa and the Middle East. This of course raises concerns regarding the effects of regional factors. We attempt to address such concerns in additional within-region (and within-country) analyses of Europe and Italy in the next two sections.

IV regressions. The literature suggests that the Christian Church's policies led to the decline of unilineal descent groups such as lineages and clans (see, e.g., Goody, 1983; Greif, 2006; Mitterauer, 2010). Based on this idea, Korotayev (2003) argues for and shows a significant negative correlation between "deep Christianization" and absence of unilineal descent groups. He only considers so-called "deep" Christianization, arguing that it took the Christian Church centuries to eradicate pre-Christian norms, values, and practices, and that one would not expect a superfluous Christianization to produce any radical changes in kinship and marriage practices and norms. Korotayev (2003) coded the variable equal to 1 if a society of the Ethnographic Atlas sample had been Christianized for at least 500 years prior to the ethnographic present as recorded in the Ethnographic Atlas. All other societies take a value of 0.[12] Following Alesina et al. (2013) and using the procedure described above, we calculate the country-level measure of deep Christianization using all ethno-linguistic groups as shown in Fig. 3 (right).

The advantage of this measure is that it doesn't distinguish between Western and Eastern Christianity. Christian ethnicities in countries with both relatively high and low corruption, such as western Europe, and Eastern Europe and Latin America respectively, take value 1 based on Korotayev (2003)'s coded data (also see Fig. C.2 in Appendix C for the ethnicity-level map). Therefore, by using this measure, we do not simply pick up the effect of high quality institutions in Western Europe. Moreover, we control for the share of Catholics and Protestants, which further addresses the concern of such confounding.

Columns (1–2) of Table 5 report the first-stage regression results. Deep Christianization (a 500-year exposure to Christianity) is associated with a 15–25% decrease in contemporary consanguinity rates in the regressions. When the measure is used as an instrument for consanguinity rates in columns (3–4), the resulting coefficients are significant. A 1 standard deviation increase in the consanguinity rate is associated with a roughly 1 standard deviation increase in corruption. The Cragg–Donald Wald first-stage F-statistic is 21.41 in the basic model.

Table C.7 of Appendix C reports the results from the full specification, additional reduced form estimates using deep Christianization, and a third specification including religion. Not surprisingly, when religion is included, the coefficients

[10] As an estimate of the geographic distribution of ethnicities across the globe, we use Language locations from WLMS (World Language Mapping System) 19th version (2017) which corresponds to the 19th edition (2016) ISO 639-3 standard, and the 16th Edition of the Ethnologue (Languages of the World). WLMS provides a shape file that divides the world into polygons indicating the locations around the world where 7650 languages are spoken. Each of these languages are matched to one of the 1291 ethnic groups included in the Ethnographic Atlas using information from the language trees of the world's languages (Hammarström and Nordhoff, 2011). For the population of each ethnic group, we use data from Gridded Population of the World (v4) which reports the world's population in 2014 for 30 arc-second by 30 arc-second grid cells.

[11] Our measure relies on an Ethnographic Atlas measure that captures a historical preference for cousin marriage, as opposed to mere permission thereof. This is because, in our view, it is the actual practice of cousin marriage and the accompanying variation in relatedness and networks of interaction that should encourage the development of norms of favoritism.

[12] Since Korotayev used an older version of the Ethnographic Atlas, we do not have deep Christianization measures for recently added ethnic groups. We treat these societies the same way we treat other ethno-linguistic groups that have missing values.

Please cite this article as: M. Akbari, D. Bahrami-Rad and E.O. Kimbrough, Kinship, fractionalization and corruption, Journal of Economic Behavior and Organization, https://doi.org/10.1016/j.jebo.2019.07.015

Table 4
Regression analysis of the relationship between lineage endogamy and corruption. Full specification reported in Table C.6 of Appendix C.

VARIABLES	(1) Basic model	(2) both Income and Latitude	(3) and Religion
Lineage endogamy	−1.022**	−0.898**	−0.702*
	(0.440)	(0.356)	(0.369)
	[0.344]	[0.257]	[0.286]
Ethnic fractionalization	0.202	0.209	0.225
	(0.479)	(0.424)	(0.412)
Latitude	0.664	0.617	0.747
	(1.716)	(1.615)	(1.570)
Log GNI per capita		0.669***	0.547**
		(0.199)	(0.215)
Protestant			0.685
			(0.468)
Catholic			0.045
			(0.260)
Muslim			−0.375
			(0.379)
Log population	−0.110	0.010	0.028
	(0.111)	(0.100)	(0.107)
Africa	−0.598*	−0.072	−0.215
	(0.348)	(0.283)	(0.286)
East Asia	−0.170	−0.077	−0.177
	(0.323)	(0.309)	(0.318)
Latin America	−0.499*	−0.446*	−0.579**
	(0.290)	(0.248)	(0.269)
Socialist legal origin	−1.478***	−1.123***	−1.025***
	(0.286)	(0.301)	(0.294)
French legal origin	0.001	0.045	0.108
	(0.172)	(0.168)	(0.176)
German legal origin	0.371	0.075	0.098
	(0.374)	(0.341)	(0.317)
Scandinavian legal origin	0.937**	1.009***	0.568
	(0.374)	(0.336)	(0.435)
Constant	4.531***	0.739	1.024
	(1.724)	(1.606)	(1.686)
Geographical variables	yes	yes	yes
Observations	126	125	124
R-squared	0.680	0.716	0.733

Robust standard errors in parentheses. Conley standard errors in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

from both stages are insignificant, most likely due to collinearity.[13] Deep Christianization is highly correlated with the set of variables for religion: percentage of Muslims, Catholics, and Protestants (Pearson's $r = -0.77$, 0.74, and 0.34 respectively, p-value < 0.01, N=64).

Nevertheless, these results do not provide decisive causal evidence, and we acknowledge the existence of other important influences from the Church on corruption via channels other than kin-based institutions. We next report analysis exploiting historical variation in consanguinity within Italy. Since Italy is an almost religiously homogeneous country, this also has the advantage of addressing concerns regarding the effect of religion in our cross-country analyses.

3.2. Within-country analysis (Italy)

To further test the hypothesis that sub-ethnic fractionalization causes corruption, while controlling for country-specific institutional and cultural factors, we collected data on corruption and consanguinity across Italian provinces. Our consanguinity data are from Cavalli-Sforza et al. (2004) and are based on records of Papal dispensations for consanguineous marriages kept in the Vatican archives. Because consanguinity was officially banned by the church, couples who wished to circumvent the ban had to request approval from their local diocese. Detailed records of these marriages were preserved by the church and compiled in province-level statistics. We report the consanguinity rate by province, measured as the share of all marriages in the province that were between first cousins, or closer relatives (e.g. uncle-niece marriages) for the period 1945–1964, since data from this period are available for all provinces and do not include sample variation due to World War

[13] A Farrar-Glauber Multicollinearity F-Test rejects the 'no multicollinearity' null hypothesis of deep Christianization and share of Muslims (p-value=0.002), Catholics (p-value=0.003), and Protestants (p-value=0.007).

Table 5
First-stage and second-stage regression analyses of the relationship between corruption and consanguinity (instrumented by deep Christianization).

	First stage		IV estimation	
VARIABLES	(1) Basic model	(2) both Income and Latitude	(3) Basic model	(4) both Income and Latitude
Deep Christianization	−0.248***	−0.147***		
	(0.052)	(0.053)		
Consanguinity			−6.666***	−8.926**
			(1.315)	(3.723)
Ethnic fractionalization	0.074	0.105	−0.008	0.331
	(0.066)	(0.063)	(0.489)	(0.858)
Latitude	−0.164	−0.236	−5.252**	−6.018*
	(0.329)	(0.303)	(2.448)	(3.096)
Log GNI per capita		−0.105**		−0.577
		(0.052)		(0.821)
Log population	0.023	0.008	−0.165	−0.198
	(0.023)	(0.020)	(0.205)	(0.240)
Africa	0.114	0.066	0.709	0.704
	(0.096)	(0.097)	(0.639)	(0.793)
East Asia	−0.189***	−0.145***	−0.907*	−1.092*
	(0.055)	(0.050)	(0.493)	(0.648)
Latin America	−0.007	−0.035	−1.435***	−1.602***
	(0.048)	(0.042)	(0.379)	(0.508)
Socialist legal origin	0.065	0.005	−1.096***	−1.280***
	(0.042)	(0.039)	(0.349)	(0.457)
French legal origin	0.010	−0.010	−0.009	−0.095
	(0.036)	(0.030)	(0.280)	(0.327)
German legal origin	−0.200***	−0.068	0.146	0.424
	(0.066)	(0.071)	(0.423)	(0.603)
Scandinavian legal origin	0.057	0.061	0.955*	1.106
	(0.058)	(0.056)	(0.544)	(0.710)
Constant	0.208	0.752**	10.276***	13.746**
	(0.318)	(0.369)	(2.082)	(5.200)
Geographical variables	yes	yes	yes	yes
Observations	64	64	64	64
R-squared	0.865	0.884	0.719	0.633

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

II.[14] The underlying data, however, date back to 1911, and Cavalli-Sforza et al. (2004) provide evidence that heterogeneity in consanguinity rates across Italy at the start of their time period was unrelated to a large vector of demographic, social and economic variables (e.g. birth/death rates, infant mortality, population density, immigration/emigration, literacy, and industrialization). Moreover, their evidence suggests that trends in consanguinity in Italy were similar across the entire country, despite substantial differences in levels.

Data on actual corruption crimes in Italy are not available for the provinces. Therefore, exploiting a known link between corruption and associative crime (i.e. criminal conspiracy and mafia association, see Fiorino et al., 2012), we use the latter as a proxy for corruption. The link is straightforward: associative crime reflects charges leveled at members of criminal organizations that seek to enrich themselves at the expense of others and of the government. Corruption is essential for such criminal organizations because it facilitates the operation of illegal markets for goods and services such as cigarettes, drugs, prostitution, gambling, as well as activities such as car-theft, extortion, tax evasion, etc. Corruption allows criminal organizations to obtain information about potential attempts to subvert their operations (by law enforcement or competitors), supports the deletion, falsification or destruction of incriminating evidence, and may be used to neutralize judges, prosecutors, police or experts who might interfere with their plans.[15,16]

[14] Due to high consanguinity rates in Sicily and Sardinia, the church relaxed its rules such that dispensations were not required for marriages between relatives more distant than first cousins in those regions, though dispensations for such marriages remained mandatory elsewhere. Thus to ensure comparability of the data across provinces, we restrict attention to first cousin marriages and closer. Note that the coefficient of relatedness between first cousins (in a randomly mating population) is $r = 1/8$ vs. $r = 1/32$ for second cousins.

[15] A model by Kugler et al. (2005) connects corruption and associative crimes, based on criminal organizations' attempts to avoid punishment by bribing law enforcement and otherwise engaging in local corruption.

[16] The proposed connection is also borne out in the available data. Data on corruption perception provided by ISTAT based on surveys in 2015–2016 exists at the region level (N = 20). Aggregating our proxy measure of corruption, "associative crime", to the region level, the two measures are highly correlated, even with a small number of observations, suggesting that associative crime is a reasonable proxy for corruption (Pearson's $r = 0.31$, p-value = 0.001). See also Gounev and Bezlov (2010) who analyzed links between organized crime and corruption for The European Commission. Their analysis includes case

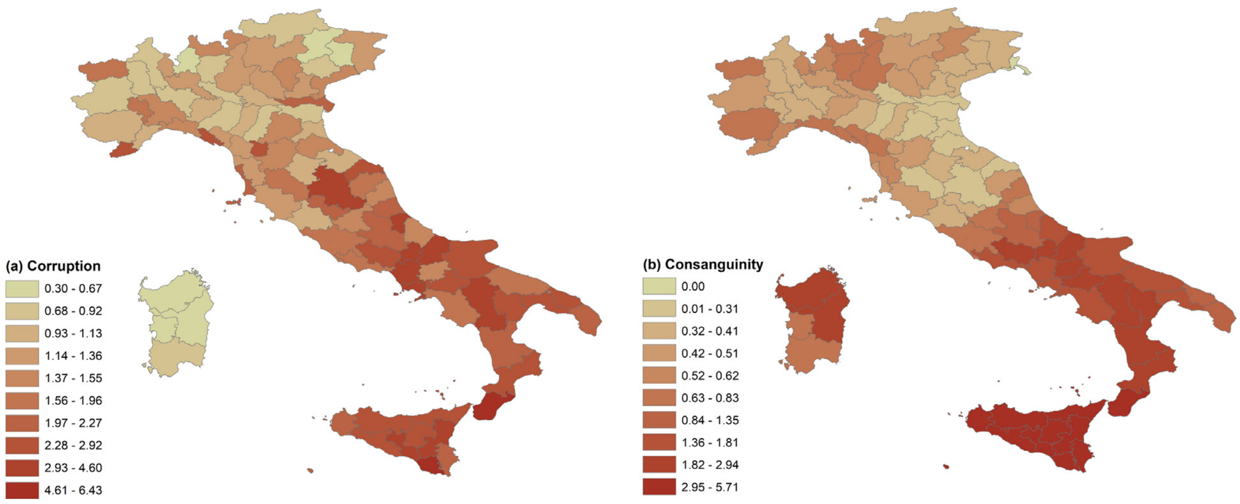

Fig. 4. Associative crimes and consanguinity in Italy.

Thus, we proxy for province-level corruption by using the number of associative crimes per 100,000 inhabitants of the province (via ISTAT). Our data on associative crimes and control variables cover the period 2000–2013. All variables are described in detail in Appendix B, Table B.2. Because of missing data on consanguinity or corruption (partly due to changes in the number of provinces since the year 2000), our analysis is based on data from 101 provinces.

Fig. 4 displays associative crimes and consanguinity by province in Italy. Overall, we find a positive and highly significant correlation between consanguinity rates and associative crimes in Italian provinces (Pearson's $r = 0.65$, p-value < 0.001, N = 101).

As in the cross-country analysis, we regress associative crimes on a number of controls and include consanguinity rates as our proxy for sub-ethnic fractionalization. In such analyses, per capita income and the relative size of the agricultural sector "are often used as proxy variables for the level of development" (Del Monte and Papagni, 2007, p. 390). Again, due to the likely endogeneity of income per capita to corruption, we prefer to use the share of agriculture in the regression, though we report specifications including both. In fact, the two variables are highly correlated (Pearson's $r = -0.39$, p-value < 0.001, N=101), reiterating that share of agriculture is a good proxy for per capita income. We also control for the population of the provinces. These variables are averaged over 2000–2013. We also include a variety of geographic controls: mean annual temperature and precipitation, soil suitability for agriculture, distance to the coast, average elevation, average slope, and ruggedness.

In addition to our correlational estimates using consanguinity rates, we also attempt to address causality by exploiting plausibly exogenous variation in historical exposure to the marriage laws of the Catholic church due to Arab, Byzantine, and Langobardic (aka Lombardic) domination of southern Italy for centuries. These historical episodes led to suppression of the Catholic church in some regions and thus facilitate an instrumental variables approach that provides some evidence of a causal relationship between consanguinity and corruption in Italy.

Note that in our cross-country analyses, we define our consanguinity variable between 0 and 1 to make its coefficients comparable with ethnic fractionalization (also defined 0–1). However, in this analysis and the analysis in the next section comparing European regions, the consanguinity rate never exceeds 6 percent (0–0.06 if defined as before), so we define it in percentage terms (0–100).

3.2.1. Main findings

Table 6 displays the results of our first set of regressions. Our specification in column (1) reveals a large and significant coefficient on agricultural share of income, indicating that less developed parts of Italy also exhibit lower institutional quality. In column (2), we introduce our proxy for sub-ethnic fractionalization: the consanguinity rate.

From the estimates in column (2), the effect of consanguinity on associative crimes is quite large: a 1 standard deviation increase in consanguinity rate is associated with a roughly 0.6 standard deviation increase in associative crimes per hundred thousand inhabitants in the province. Moreover, the difference in consanguinity between the least and the most consanguineous province is associated with an increase of roughly 2.7 associative crimes per hundred thousand population, about 44% of the difference between the most and least-corrupt region. Columns (3) shows that our results are robust when we also control for income. Contrary to the cross-country analysis, population, income and share of agriculture are not significant in the regressions. This suggests that between-province variability of population or the degree of development is

studies on several European countries including Italy, where corruption and organized crime "are closely intertwined" (p. 157) and "criminal organizations such as the mafia are the most visible in terms of exercising power" (p. 163).

Please cite this article as: M. Akbari, D. Bahrami-Rad and E.O. Kimbrough, Kinship, fractionalization and corruption, Journal of Economic Behavior and Organization, https://doi.org/10.1016/j.jebo.2019.07.015

Table 6
Regression analysis of the relationship between consanguinity and associative crimes in Italy.

VARIABLES	(1)	(2) Basic model	(3) both Income and Share of agriculture
Consanguinity		0.582**	0.584**
		(0.211)	(0.214)
		[0.172]	[0.175]
Share of agriculture	2.943	1.589	1.950
	(3.133)	(3.047)	(3.168)
Log value added per capita			0.024
			(0.108)
Log population	0.117	−0.018	−0.001
	(0.112)	(0.092)	(0.126)
Mean Temperature	0.039	−0.044	−0.044
	(0.060)	(0.062)	(0.063)
Mean precipitation	9.710***	−0.902	−0.955
	(3.361)	(3.573)	(3.619)
Suitability for agriculture	0.495	0.097	0.083
	(0.526)	(0.405)	(0.428)
Distance to coast	−0.007***	−0.006***	−0.006***
	(0.002)	(0.002)	(0.002)
Elevation	0.094	0.094	0.094
	(0.071)	(0.074)	(0.075)
Ruggedness	−0.057	−0.205	−0.204
	(0.134)	(0.140)	(0.141)
Constant	−0.550	1.668	1.536
	(1.544)	(1.243)	(1.403)
Observations	101	101	101
R-squared	0.345	0.475	0.475

Robust standard errors, clustered at region level, in parentheses. Conley standard errors in brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

too low to capture its effect on corruption. Consanguinity is associated with significantly higher associative crimes in all specifications.

In our primary analyses, we cluster standard errors at the region level. The results are robust to using Conley (1999) standard errors to account for spatial dependence with different cutoffs. For example, in the table, the Conley standard errors of consanguinity with cutoff of 3 decimal degrees is reported in the brackets. The results also hold in spatial autoregressive (SAR) model estimates.[17]

Robustness checks. There are a few other potential confounds which, if addressed, would increase confidence in our results. In Table D.1 of Appendix D, we report further analyses building on the basic model and including additional variables that have been proposed to mediate institutional quality (hence corruption) in the literature; these include differences between the south and north in Italy (Banfield, 1958; Putnam et al., 1994; Bigoni et al., 2016), measures of civil society (Putnam et al., 1994), family types (Todd, 1990), family ties (Alesina and Giuliano, 2014; Alesina et al., 2015), and historical political domination by various groups (Putnam et al., 1994; Di Liberto and Sideri, 2015). Our regression analyses show that consanguinity remains significant at the 5% level with any combinations of potential confounding factors, except when latitude is included in regressions. See Table B.2 of Appendix B for description of variables used for the analyses.

Our results are also robust to altering cutoff dates for computing consanguinity rates from 1945 to 1964 to 1950–1964, 1955–1964, or 1960–1964. Table D.2 of Appendix D shows results for the baseline regression.

Since associative crime is not a measure of actual corruption, but only of the crimes reported and detected by the police, it may underestimate the true phenomenon because of judicial inefficiency. Moreover, the most corrupt regions may be those in which such crimes are least likely to be reported or detected, which further suggests that we may underestimate the relationship between consanguinity and corruption. Thus, in Appendix D, we show that our regression results are robust to using an alternative measure of corruption from Golden and Picci (2005) based on computing the ratio of the value of existing physical infrastructure stocks (in 1998) to the expected value of infrastructure given government expenditures over the period 1954–1997. “The intuition underlying this measure is that, all else equal, governments that do not get what they pay for are those whose bureaucrats and politicians are siphoning off more public monies in corrupt transactions” (Golden and Picci, 2005, p. 41). The measure is available for 90 provinces for which we also have consanguinity data, and it is also highly correlated with consanguinity (Pearson’s $r = -0.52$, p-value < 0.001, $N = 90$). See Fig. D.1 and Table D.3 of Appendix D.

[17] For example, allowing for spatial lags of the dependent variable, independent variable, and error in the basic model of the Table 6, the estimated ‘direct effect’ of consanguinity in a province has nearly half of the magnitude of our baseline estimate: 0.305 (p-value=0.065), and we see a large spillover effect to other provinces: 2.694 (p-value=0.088). Full estimates available upon request.

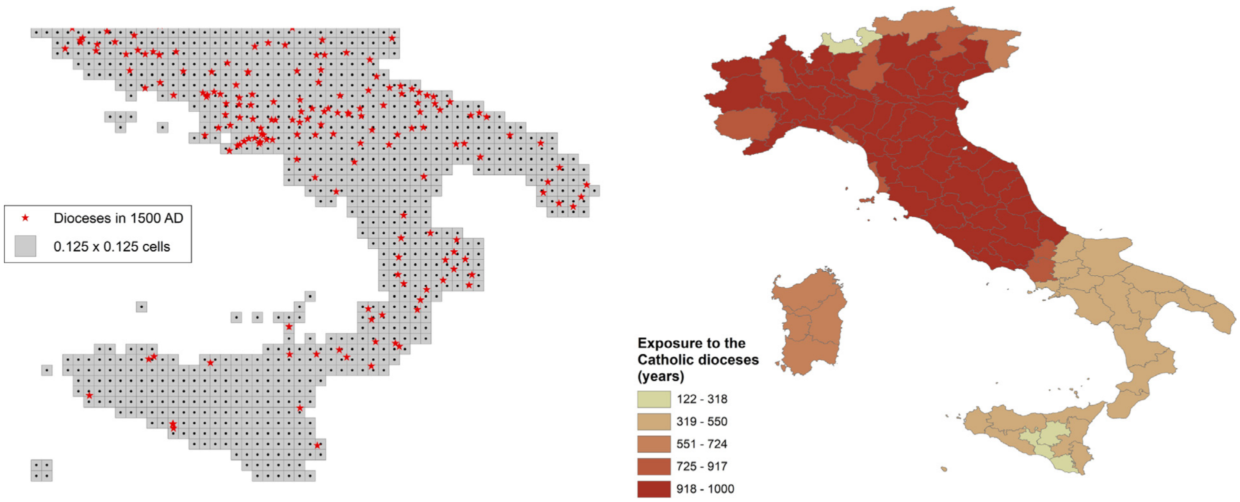

Fig. 5. (Left) A sample of cells and locations of dioceses in 1500 CE used to calculate the measure for exposure to the Catholic dioceses. (Right) Exposure to the Catholic dioceses between 500 and 1500 CE across Italian provinces.

IV Regressions. Some of the modern cultural differences between northern and southern Italy might be driven, in part, by historical differences in marriage patterns between the two regions, perhaps “as a remote consequence of Arab domination in Sicily and southern Italy in the eighth to the eleventh centuries” (Cavalli-Sforza et al., 2004, p. 3). Moreover, southern Italy was governed by Byzantium and the Langobardic kingdoms for centuries—rather than the Catholic Church—where Bishops exercised almost no jurisdictional authority, and local customs prevailed (Schulz et al., 2018). The possibility of differences in mating patterns due to a historical event such as Arab domination provides a potential exogenous variation that is worthwhile to explore. These events introduced exogenous variation in the intensity of exposure to the Catholic church (and its policies on legitimate marriage). To capture this variation, we use data and methodology from Schulz et al. (2018).

The data on location and dates of establishment and suppression of Catholic dioceses is available for Italy at half-centuries between 500 and 1500 CE. To calculate Catholic Church exposure at the province level, we first divided Italy into cells of 0.125 $\times$ 0.125 decimal degrees (about 14 by 14km at the equator). Then for each cell in each half-century, the distance from the centroid of the cell to the nearest diocese is calculated (see Fig. 5, left). A cell takes value of 1 if there was a diocese within 50km of the cell’s centroid, and takes value 0 otherwise. Finally, for each cell, exposure to the Catholic dioceses over all half-centuries are added, yielding the duration of the cell’s exposure to the Catholic dioceses in years (see Fig. 5, right).

Table 7 reports reduced form estimates alongside our IV estimates (and the associated first-stage from the IV) using the measure of Church exposure as an instrumental variable for consanguinity. Reduced form regression results in columns (1–2) show that provinces with higher medieval exposure to the Catholic dioceses exhibit lower rates of associative crime today. The first-stage results in columns (3–4) show that higher medieval exposure of a province to Catholic dioceses is associated with 1% lower cousin marriage rate in the province today. The Cragg-Donald Wald first stage F statistic is 17.51 in the basic model. The IV results in columns (5–6) show that provinces with longer exposure to the Catholic dioceses exhibit lower rates of associative crime. A 1 standard deviation increase in consanguinity rate increases associative crimes per hundred thousand inhabitants by almost 1 standard deviation.[18]

Since exposure to the Catholic dioceses pre-dates cousin marriages during the 20th century, there is little concern about reverse causality. However, again, we cannot rule out the impact of historical exposure to dioceses on today’s corruption level through other channels such as education, literacy, religiosity, post-war male population, etc.

Table D.4 of Appendix D reports reduced form and instrumental variable estimations using the alternative measure of corruption from Golden and Picci (2005), where exposure to the Church and consanguinity are again highly significant.

In sum, in our baseline specifications and controlling for a variety of possible climatic, geographical, and historical confounds, consanguinity is a significant predictor of corruption in Italy. Moreover, our IV results provide some evidence that the relationship may be causal. As in the cross-country analysis, these data are highly consistent with a model in which corruption is in part driven by sub-ethnic fractionalization, as reflected in consanguinity.[19]

[18] Latitude and the South dummy variable are highly correlated with the measure of exposure to Catholic dioceses (Pearson’s $r = 0.67$ and -0.73 respectively, p-value < 0.000, N=101). When they are included in the regression, consanguinity becomes insignificant. This is not surprising since the differences in the exposure to the Catholic dioceses in Italy are mediated by differences in latitude; the Church’s influence was weaker in southern regions because of its geographic distance from Rome and due to historical Arab, Norman, and Lombard domination of these regions.

[19] One further implication of our theory that our current data does not allow us to test is that *local crime*, in which neighbors are targets, should be less common in regions with high consanguinity rates. Buonanno and Vanin (2017) provide related evidence consistent with our findings and with this

Table 7
Exposure to the Catholic dioceses (units in centuries) as an instrument for consanguinity.

	Reduced form		First stage		IV estimation	
VARIABLES	(1) Basic model	(2) both Income and Share of agriculture	(3) Basic model	(4) both Income and Share of agriculture	(5) Basic model	(6) both Income and Share of agriculture
Exposure to the Church	−0.134**	−0.135**	−0.169***	−0.171***		
	(0.061)	(0.061)	(0.048)	(0.049)		
Consanguinity					0.795***	0.791***
					(0.197)	(0.213)
Share of agriculture	1.569	0.814	0.598	−1.252	1.093	1.804
	(3.348)	(3.708)	(3.020)	(3.191)	(2.689)	(3.200)
Log value added per capita		−0.049		−0.119		0.046
		(0.120)		(0.082)		(0.107)
Log population	0.093	0.057	0.202**	0.114	−0.068	−0.033
	(0.117)	(0.166)	(0.086)	(0.115)	(0.132)	(0.147)
Mean Temperature	0.003	0.003	0.096***	0.096***	−0.074	−0.074
	(0.061)	(0.061)	(0.035)	(0.035)	(0.054)	(0.058)
Mean precipitation	6.143	6.134	13.755***	13.735***	−4.789	−4.726
	(4.429)	(4.420)	(2.953)	(2.971)	(5.667)	(6.147)
Suitability for agriculture	0.958	0.985*	1.266**	1.332***	−0.048	−0.069
	(0.557)	(0.556)	(0.516)	(0.502)	(0.462)	(0.520)
Distance to coast	−0.007***	−0.007***	−0.001	−0.001	−0.006***	−0.006***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Elevation	0.090	0.089	−0.005	−0.006	0.094	0.095
	(0.068)	(0.067)	(0.055)	(0.055)	(0.075)	(0.082)
Ruggedness	−0.073	−0.076	0.236**	0.226*	−0.260	−0.255
	(0.147)	(0.153)	(0.114)	(0.117)	(0.207)	(0.219)
Constant	0.798	1.094	−2.117	−1.391	2.480	2.194
	(2.011)	(2.305)	(1.378)	(1.609)	(1.976)	(2.054)
Observations	101	101	101	101	101	101
R-squared	0.384	0.385	0.776	0.780	0.457	0.459

Robust standard errors, clustered at region level, in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

3.3. Consanguinity and corruption across European regions

Data designed to capture corruption alone did not exist for European regions until the creation of the Quality of Governance (QoG) Data by the EU Commission's funded projects (Charron et al., 2014; 2015). QoG includes data on the European Quality of Government Index (EQI) which is the result of novel survey data (years 2010, 2013, and 2017) on region-level (unfortunately not province-level) governance across the EU. In each round of the survey, individuals (450 per region in 2013, for example) are asked questions about the extent to which they perceive and experience corruption, and about the quality and impartiality in provision of services such as education, healthcare services, and law enforcement. The EQI of a region is the aggregation of the scores for the three pillars: corruption, quality, and impartiality.[20]

The corruption pillar is derived from questions about perceived corruption of public education/public health care/law enforcement in the respondent's region, the amount of perceived bribery in the respondent's region, and the respondent's own experience with bribery in public education/public health care/law enforcement in the respondent's region. We use the normalized version (0–100) of the data as our dependent variable, and we again assess the association between sub-ethnic fractionalization and corruption, using consanguinity as a proxy for sub-ethnic fractionalization.[21]

Within-country cousin marriage data in Europe is available for Italy, France, Spain, and Turkey (Schulz et al., 2018). Unfortunately, EQI data is not available for Turkey. Therefore, we end up with a sample of 57 regions with available data on both the EQI corruption index and cousin marriage: 21 in Italy[22], 21 in France, and 15 in Spain.

Despite the low number of observations, our regression analyses reveal a robust correlation between consanguinity and corruption across the 57 European regions, all of which are historically Catholic. Table 8 reports our regression analyses for the 57 regions under a variety of specifications. In the baseline model we include the same vector of geographic controls (temperature, precipitation, suitability for agriculture, distance to coast, elevation, and ruggedness), latitude, population, and country fixed effects. In the baseline specification, a 1% increase in consanguinity (95% of a standard deviation) is associated with a 5.2 unit increase in the corruption index (33% of a standard deviation).

hypothesis as well. Using the distribution of surname frequencies from Italian phonebooks, they show that municipalities with more surname concentration (i.e. with more in-marriage) exhibit both more tax evasion (of a federal tax) and less local crime.

[20] The data is available online at https://qog.pol.gu.se/data/datadownloads/qog-eqi-data.

[21] The measure is highly correlated (Pearson's $r=-0.65$, p-value=0.002, N=20) with the corruption perception index across Italian regions provided by ISTAT based on surveys in 2015–2016.

[22] The Provinces of Trento and Bolzano are treated as separate regions in these data.

Please cite this article as: M. Akbari, D. Bahrami-Rad and E.O. Kimbrough, Kinship, fractionalization and corruption, Journal of Economic Behavior and Organization, https://doi.org/10.1016/j.jebo.2019.07.015

Table 8
Regression analyses of the relationship between consanguinity and corruption in European regions (Italy, France, Spain). Lower values represent higher corruption.

VARIABLES	(1) Consanguinity	(2) and baseline controls	(3) and fixed effects (Basic model)	(4) and income
Consanguinity	−9.220***	−7.449***	−5.198***	−3.550**
	(1.722)	(1.955)	(1.536)	(1.319)
	[2.681]	[2.120]	[1.377]	[1.217]
Log population		−1.456	−0.957	−2.389**
		(1.382)	(0.836)	(0.955)
Log GDP per capita				19.186***
				(6.571)
Latitude		2.697***	1.125	0.135
		(0.785)	(0.722)	(0.693)
Mean temperature		1.784**	−0.444	−0.534
		(0.878)	(0.600)	(0.577)
Mean precipitation		0.182**	0.078	0.077*
		(0.084)	(0.047)	(0.043)
Suitability for agriculture		−0.007*	−0.003	−0.006**
		(0.004)	(0.003)	(0.003)
Distance to coast		0.000	−0.000	−0.000
		(0.000)	(0.000)	(0.000)
Elevation		0.017*	−0.003	−0.009
		(0.009)	(0.006)	(0.007)
Ruggedness		−0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Constant	61.830***	−41.578	32.093	−92.207
	(2.219)	(47.008)	(36.694)	(59.236)
Country FEs			yes	yes
Observations	57	57	57	57
R-squared	0.316	0.709	0.883	0.906

Robust standard errors in parentheses. Conley standard errors in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The result is robust to controlling for income per capita of regions and adjusting standard errors for spatial dependence with different cutoffs. For example, Conley standard errors for the coefficient on consanguinity using a cutoff of 5 decimal degrees are reported in brackets. The result also holds when we estimate a spatial autoregressive (SAR) model.[23]

Robustness checks. Quality of Governance (QoG) includes data on European regions that allow us to run analyses controlling for additional confounding factors, parallel to our discussion and analyses of Italian regions. Table E.1 show that results are robust to controlling for generalized trust, percentage of Protestants, fractionalization due to the presence of political parties, and civic engagement, as proxied by voter turnout. The results are presented in Table E.1 of Appendix E.

We also applied our instrumental variable approach using the measure of historical exposure to Catholic dioceses. Considering the low number of observations and inclusion of country fixed effects in our basic model, we are able to get significant results only when we include a subset of geographic controls (excluding temperature, precipitation, and latitude) in the basic model. Table E.2 of Appendix E provides the results.

4. Conclusion

Countries around the world exhibit vast differences in levels of corruption, and understanding the sources of these differences is crucial to improving governance. We provide evidence from cross-country and within-country regression analyses that in-marriage and concomitant *sub-ethnic fractionalization* is an important determinant of corruption. In regions with more in-marriage practices in the form of consanguinity, corruption is relatively prevalent, even after controlling for previously studied deep determinants of corruption. We argue that in-marriage practices, which increase the intensity of ties to the in-group and increase relatedness between individuals at the local level, raise the relative returns to norms of kin altruism, nepotism, and favoritism; while, out-marriage practices raise the relative returns to norms of impartial, generalized cooperation. While we focus on the example of cousin marriage, we would expect the same logic to apply to other local in-marriage practices (e.g. locally endogamous marriage within a caste). Our findings suggest that historical differences in mating practices (due to religion, geography, and local circumstance) may have had a powerful influence on today's norms. Although cross-country and within-country analyses using exposure to Christianity reveal statistical patterns that are consistent with our theory, we should note that such analyses are susceptible to potential confounds through institutional or

[23] Allowing for spatial lags of the dependent variable, independent variable, and error in the basic model of the Table 8, the estimated direct effect of consanguinity in a region has a comparable magnitude to our baseline specification: -3.207 (p-value=0.008), and we see a large spillover effect to other regions: -25.915 (p-value=0.034). Full estimates available upon request.

Table A1
Genetic and religious classification of consanguinity (Bittles 2012, p. 16).

Biological relationship	Genetic relationship	Roman classification	Germanic classification
First cousin	Third degree	Fourth degree	Second degree
Second cousin	Fifth degree	Sixth degree	Third degree
Third cousin	Seventh degree	Eighth degree	Fourth degree

human capital channels. Future work should continue to assess the extent to which the robust relationship we document is causal.

Appendices to Akbari, Bahrami-Rad and Kimbrough (2019)

Appendix A. Religion and Consanguinity

As noted in the text and in Table 1, religious traditions have diverse relationships to cousin marriage practices. Below we summarize Christian and Islamic views on consanguinity.

Catholic and Orthodox Christianity. In the early Christian era, consanguineous marriages were common,[24] and there are few direct biblical prohibitions on marriage among close kin.[25] The laws and customs of the Roman empire with respect to domestic life, "conformed to patterns that were wide-spread throughout Mediterranean and Middle East" that permitted and even encouraged consanguinity. Thus the Christian religion emerged in a setting in which such practices were tolerated. In other regions to which Christianity later spread such as Greece and Egypt, the earlier presence of close marriage was yet more marked (Goody, 1983, p. 39).

However, in the centuries after the conversion of the Roman Empire to Christianity, radical changes occurred with respect to the issue of marriage to kin. For the first two-hundred years after Constantine, the legality of consanguinity was in flux.[26] Roman Catholic authority stepped in to settle the issue. Pope Gregory I (CE 540–604) in a letter to Augustine, 'the first Bishop and apostle of the English' confirms that a certain secular law in the Roman state allows cousin marriage and advises that 'it is necessary that the faithful should only marry relations three or four times removed, while those twice removed must not marry in any case' (Goody, 1983, p. 36). This letter was significant because through it, the Church asserted (implicitly) its jurisdiction over marriage and the family, a position that Christian churches still maintain today.

Within the Roman Catholic Church, the stringency of the restrictions continued to increase over time. A canon attributed to various popes and embodied in a letter of Pope Gregory III (CE 731–741) forbade marriage to the seventh degree of consanguinity in CE 732 (i.e. to 3rd cousins, see Table A.1), and confusion resulting from differences in the Germanic and Roman systems of consanguinity evaluation eventually led to the ban being extended all the way to 6th cousins in 1076 (7th degree in the Germanic system). Not only were these extended prohibitions attached to blood ties, but they were also assigned to affinal kinship such as marriage to the dead brother's widow, spiritual kinship such as marriage to godchildren and fictional kinship such as adoption, "producing a vast range of people, often resident in the same locality, that were forbidden to marry" (Goody, 1983, p. 56). While the bans were later loosened and dispensations have typically been available in specific cases, prohibitions on first cousin marriage remain in effect today.[27] Similarly, within the Greek Orthodox Church, first-cousin marriages were prohibited by CE 692, a policy which remains in place.

It is worthwhile to briefly discuss the possible reasons behind Church prohibitions on consanguineous marriages. From Gregory's letter to Augustine, it is clear that he understood the potential negative health consequences of inbreeding - 'the

[24] On one hand, it has been suggested that with the aim of favoring outbreeding, pre-Christian Roman law forbade all unions among people within the seventh degree of consanguinity, and the Church "initially followed the Roman" regulations on consanguineous marriages (Cavalli-Sforza et al., 2004, p. 29,31). On the other hand, some authors argue that there is not enough evidence suggesting prohibition of consanguineous unions at the very earliest stage in Roman history (Goody, 1983, p. 53). However, it seems that in the early Christian era, consanguineous marriages were common either due to "relaxed" (Cavalli-Sforza et al., 2004, p. 29) earlier prohibitions or because Roman "law had nothing to say against most forms of close marriage" (Goody, 1983, p. 39). For example, the first Christian emperor, Constantine the Great married his son and daughters to his half-brother's children (Goody, 1983, p. 53).

[25] The only types of forbidden relationships in the Bible are found in the Levitical prohibitions in The Old Testament. For example, a man could not marry his mother (Lev. 18:7), his sister (Lev. 18:12, 20:19), his father's sister (Lev. 18:12, 20:19), his mother's sister (Lev. 18:13, 20:19), or his son's daughter (Lev. 18:10). A man could marry his niece (Cavalli-Sforza et al., 2004, p. 29).

[26] Following the acceptance of Christianity as the official religion of the Eastern Roman Empire, Theodosius the Great (AD 378–395) condemned unions between first cousins in a law made in AD 384, although it was still possible to effect such marriages by imperial dispensation. In AD 405, his son, Arcadious (AD 378–408) legalized cousin marriages once again for the Eastern Roman Empire, however, his younger brother and the Western Roman emperor, Honorious (384–423), who himself married his dead wife's sister, permitted marriages among cousins in AD 409 only if the parties obtained an imperial dispensation. It was not until later that under civil legislation they became freely permissible once again. In AD 533, the validity of cousin marriage in the secular law of the Eastern Roman Empire, the *institutes* of Justinian (482–565), was recognized as perfectly legitimate (Goody, 1983, p. 55; Cavalli-Sforza et al., 2004, p. 29–30; and Bittles, 2012, p. 16).

[27] The prohibition on consanguineous marriages was reduced to 4th degree relationships (third cousin) in AD 1215. The ban also reduced to 2nd degree relationships (first cousin or closer) for South American Amerindians in 1537, later for the indigenous population of The Philippines, and for black populations in 1897. After 1917, the consanguinity prohibition was reduced in all populations, initially to second cousins or closer and in 1983 to first cousins or closer, which remains current law (Cavalli-Sforza et al., 2004; Bittles, 2012; Goody, 1983).

Please cite this article as: M. Akbari, D. Bahrami-Rad and E.O. Kimbrough, Kinship, fractionalization and corruption, Journal of Economic Behavior and Organization, https://doi.org/10.1016/j.jebo.2019.07.015

offspring of such marriages cannot thrive' - and that he had a moral opposition to incest, derived from his reading of scripture (Goody, 1983, p. 36–37). More importantly for our purposes, it seems that the potential impact on social behavior had not gone unnoticed by the Church. St. Augustine of Hippo, writing in the early 5th century CE, noted that restrictions on consanguinity expand the scope for mutually beneficial cooperation: "for affection is now given its proper place, so that men, for whom it is beneficial to live together in honorable concord, may be joined to one another by the bonds of diverse relationships: not that one man should combine many relationships in his sole person, *but that those relationships should be distributed among individuals, and should thereby bind social life more effectively by involving a greater number of persons in them*" (Augustine, 1998, Book 15, Ch. 16, *italics* added). Similar arguments were later echoed by Thomas Aquinas, who worried that "consanguineous marriages would 'prevent people widening their circle of friends'" (Goody, 1983, p. 57).[28] This suggests an awareness of the sorts of favoritism that might generate corruption.

It seems that as the Church extended prohibitions to remoter degrees of consanguineous marriage, large kinship groups such as clans, lineages, and tribes gradually disappeared throughout Europe. There is a large and significant correlation between the spread of Christianity (for at least 500 years) and the absence of clans and lineage groups in Europe (Greif, 2006, p. 309) which may have facilitated the emergence of impartial economic and political institutions, nuclear families and individualistic culture.

Protestant and Anglican Christianity. As the Protestant Reformation emphasized a return to scripture, Martin Luther's (1483–1546) views on consanguinity were based on the Levitical prohibitions and did not include a ban on cousin marriages; similarly, John Calvin (1509–1564) and his followers based their views on scripture, though they extended Levitical guidelines to a wife's relatives so that affinity and consanguinity were treated equivalently. The Church of England grew directly out of a dispute between Henry VIII (1491–1547) and the Catholic Church about marriage law, and its creation led to changes in consanguinity law that allowed for first cousin marriage (Goody, 1983, p. 168–177).

In the emergent Protestant denominations, marriages up to and including first-cousin unions were permitted under Reformed Church law.[29] However, "many Protestant reformers discouraged marriages within the third degree. While Luther did not think they were positively harmful, he considered them to be 'inexpedient on the ground that people would marry without love merely to keep property within family, while poor women would be left spinsters'" (Goody, 1983, p. 181-2). In practice, consanguinity rates remained low in the Protestant world, as seen in Fig. 2 in Section 3 above, perhaps due to ingrained norms from centuries of prohibition.

Islam. Today, consanguinity rates are high in Islamic countries, though "contrary to widespread Western opinion, there is no specific guidance in the Qur'an to encourage consanguinity" (Bittles, 2012, p. 22) and "marriage between cousins is not prescribed by the Qur'an" (Courbage and Todd, 2014, p. 33). The permitted degrees of consanguinity within Islam closely match the Levitical guidelines with one exception: a prohibition of uncle-niece marriage (Qur'an 4:23). In addition to the Qur'an, Muslims recognize two other sources of Islamic Law (Sharia) which bear on consanguinity: the Hadith (oral pronouncements of the Prophet Mohammad) and the Sunnah (the deeds of the Prophet).[30] "The overall attitude to consanguineous marriage within Islam is somewhat ambiguous" (Bittles, 2012, p. 22) since a Hadith of the Prophet Muhammad stated: "Do not marry cousins as the offspring may be disabled at birth" (Akrami and Osati, 2007, p. 314); while, according to the Sunnah, two of Muhammad's wives were his first cousins, and the Prophet Muhammad married his daughter, Fatima to his cousin Ali.[31] "Thus, [despite the content of the Hadith] for Muslims, the practice of cousin marriage could potentially be interpreted as following the example provided by the Sunnah" (Bittles, 2012, p. 22).

In any case, evidence suggests that cousin marriage customs in Islamic countries "probably antedated the spread of Arabs" (Cavalli-Sforza et al., 2004, p. 284) and "predated Islam" (Courbage and Todd, 2014, p. 33). A preference for consanguinity is reflected in "the well known Iranian proverb 'the first cousin's marriage contract has been recorded in heaven,'" though the preference "is merely a cultural and local custom rather than a religious belief" (Akrami and Osati, 2007, p. 315). Nevertheless, cousin marriage is not forbidden or clearly discouraged in Islam, and it is likely that "Islam served as a vehicle for the geographical spread of the endogamous practices [...] the model was adopted not for religious reasons but because it was the practice of a prestigious group, the Arabs, bearers of the message of the Qur'an" (Courbage and Todd, 2014, p. 33-4). One explanation suggesting why the conversion to Islam may have encouraged the high prevalence of cousin marriages in Islamic countries is the Quranic law on the inheritance of property which entitled daughters to inherit half of the amount received by sons, and wives to receive a share from their husbands. A dowry (Mahr) also is specified in Islamic law as part of the marriage arrangement. Under these circumstances, consanguineous marriages could prevent part of the family wealth from leaving the clan (Goody, 1983, p. 32).

Other religions. In Table 1, we summarized other religions' views on consanguineous marriage based on Bittles (2012), p. 23–28. Suffice to say that there has historically been a great diversity of practices around the world, with many groups

[28] Goody (1983) also suggests a number of other possible, complementary reasons for the bans, including attempts to increase fealty to the Church by diminishing the role of the family (rooted in Jesus' entreaties to deny the family, e.g. Matthew 10: 34–37) and Church desire to acquire property by leaving the deceased without eligible heirs. The latter argument finds support in the tendency of the Church to oppose any practice that expanded the number of kin, e.g. divorce, polygamy, and adoption.

[29] "An exception to this generalization is provided by the State Lutheran Church of Sweden, which until 1680 refused to recognize first-cousin unions, and from 1680 to 1844 approval was required from the King of Sweden before a first-cousin union could proceed" (Bittles, 2012, p. 20).

[30] In Shi'a Islam, law also derives from the oral pronouncements and deeds of the twelve Imams.

[31] The first Imam, in the Shi'a tradition, and the fourth of the Rashidun Caliphate according to Sunni tradition.

banning cousin marriage outright and certain groups favoring even closer marriages (e.g. Dravidian Hindu and Sephardic Jewish preference for uncle-niece marriages).

Appendix B. Empirical data description

Table B1
Description of the data for cross-country analysis.

Cross country regressions: Variable name	Description and source
Consanguinity	As a working definition, unions contracted between persons biologically related closer than second cousins are categorized as consanguineous. Bittles and Black (2015) provided a compilation of the proportion of consanguineous marriages from 262 journal papers and book chapters which report consanguinity percentages of 448 samples of different sizes in different locations of 72 countries based on household surveys, Roman Catholic Church dispensations, parish records, civil registrations, marriage registrations and surveys on blood donors, obstetric inpatients, hospital outpatients, hospital births, etc that include information for around 9.8 million marriages. The sample of countries is non-random since the data were collected for other purposes (e.g. the public health studies disproportionately sample high consanguinity societies since these are the societies with higher rates of genetic disorders), but the coverage is broad. Data collection periods vary from 1922 to 2013, with around 90% after 1950, 75% after 1960, 50% after 1970, 40% after 1980, 30% after 1990 and 15% after 2000. We collected this data from www.consang.net in December 2015 and computed the mean percentage of consanguineous marriages for each country by weighting the individual estimates according to sample size. Note that (i) the data on Czechoslovakia in the period 1961-64 is used for both the Czech Republic and Slovakia; (ii) some sample studies were reported twice, once at the city level and once for the province, in which case we considered only the province level report; (iii) we also ignored the studies listed as “Minorities and Isolates” to avoid overweighting outliers. Source: *Bittles and Black (2015)*.
Corruption	Measures financial corruption in the form of demands for special payments and bribes connected with import and export licenses, exchange controls, tax assessments, police protection, or loans, and actual or potential corruption in the form of excessive patronage, nepotism, job reservations, ‘favor-for-favors’, secret party funding, and suspiciously close ties between politics and business. In our tables, this measure is averaged over 1984–2011. La Porta et al. (1999) and Alesina et al. (2003) used the average of the months of April and October in the monthly index between 1982 and 1995, and we did so in the replication of their tables. Source: *Political Risk Services, International Country Risk Guide (2012).*
Deep Christianization	The fraction of the population of a country that had been Christian for at least 500 years in the pre-industrial era. The variable “deep Christianization? coded by Korotayev (2003) for the Ethnographic Atlas is a dummy variable which takes value 1 if the ethno-linguistic group in the Ethnographic Atlas had been Christianized for at least 500 years prior to the ethnographic present as recorded in the Ethnographic Atlas. Otherwise, it takes value 0. Following Alesina et al. (2013)?s methodology, the country-level measure is created by taking a population weighted average of the dummy variable for all ethnic groups living within a country. Source: *Korotayev (2003)*.
Ethnic fractionalization	Measures ethnic fractionalization; the probability that two randomly selected individuals from a population belong to different ethnic groups. Source: Alesina et al. (2003).
Ethonolinguistic fractionalization	Average value of five different indices of ethonolinguistic fractionalization. Its value ranges from 0 to 1. The five component indices are: (1) index of ethonolinguistic fractionalization 1960, which measures the probability that two randomly selected people from a given country will not belong to the same ethnolinguistic group (the index is based on the number and size of population groups as distinguished by their ethnic and linguistic status); (2) probability of two randomly selected individuals speaking different languages;(3) probability of two randomly selected individuals do not speak the same language; (4) percent of the population not speaking the official language; and (5) percent of the population not speaking the most widely used language. The data is collected using Easterly and Levine (1997). The sources of the components of the average index are (1) Bruk (1964); (2) Muller (1964); (3) Roberts (1962); (4) and (5) Gunnemark (1991). Source: La Porta et al. (1999).
Family ties	Measures the strength of family ties by looking at three variables from the World Value Survey (WVS) and the European Social Survey (EVS) which capture beliefs regarding the importance of the family in the respondent’s life, the duties and responsibilities of parents and children, and the love and respect for one’s own parents. The first question asks how important the family is in one person’s life and can be answered with (i) Very important; (ii) Rather important; (iii) Not very important; (iv) Not at all important, which in our measure of family ties, take values of 4 to 1, respectively. The second question asks whether the respondent agrees with one of the two statements (taking the values of 2 and 1, respectively): (i) Regardless of what the qualities and faults of one’s parents are, one must always love and respect them; (ii) One does not have the duty to respect and love parents who have not earned it. The third question prompts respondents to agree with one of the following statements (again taking the values of 2 and 1, respectively): (i) It is the parents’ duty to do their best for their children even at the expense of their own well-being; (ii) Parents have a life of their own and should not be asked to sacrifice their own well-being for the sake of their children. Following Alesina and Giuliano (2010), we extract the first principal component from the whole data set with all individual responses for the original variables. Source: *WVS (Six waves, 1981–2014) and EVS (four waves, 1981–2008).*

(*continued on next page*)

Table B1 (*continued*)

Cross country regressions: Variable name	Description and source
Generalized trust	Measures generalized trust by considering the following question from the World Value Survey (WVS) and the European Social Survey (EVS): "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?" The answer could be either "Most people can be trusted" or "Can't be too careful", which in our measure of trust, take values of 2 and 1, respectively. Source: *WVS (Six waves, 1981–2014) and EVS (four waves, 1981–2008)*.
Lineage endogamy	The fraction of the population of a country that preferentially favored parallel cousin marriage. Based on variable 26 of the Ethnographic Atlas, a dummy variable is created which takes value 1 if the ethno-linguistic group in the Ethnographic Atlas practices parallel cousin marriage, and takes value 0 otherwise. Following Alesina et al. (2013)'s methodology, the country-level measure is created by population weighted average of the dummy variable for all ethnic groups living within a country.
Predicted genetic diversity	The expected heterozygosity (genetic diversity) of a given country as predicted by (the extended sample definition of) migratory distance from East Africa (i.e., Addis Ababa, Ethiopia). This measure is calculated by applying the regression coefficients obtained from regressing expected heterozygosity on migratory distance at the ethnic group level, using the worldwide sample of 53 ethnic groups from the HGDP-CEPH Human Genome Diversity Cell Line Panel. The expected heterozygosities and geographical coordinates of the ethnic groups are from Ramachandran et al. (2005). Expected heterozygosities are constructed by measuring actual heterozygosity within an ethnic group at a sample of selectively-neutral loci and averaging over the loci. Source: *Ashraf and Galor (2013)*.
Geographical controls	(i) Ruggedness Terrain ruggedness measures small-scale terrain irregularities. The ruggedness calculation takes a point on the earth's surface and calculates the difference in elevation between this point and each of the points on the grid 30 arc-seconds (926 m on a meridian) in each of the eight major directions of the compass (north, northeast, east, southeast, south, southwest, west, and northwest). The terrain ruggedness index at the central point is given by the square root of the sum of the squared differences in elevation between the central point and the eight adjacent points. Then by averaging across all grid cells in the country not covered by water, each cell weighed by its latitude-varying sea-level surface, the average terrain ruggedness of the country's land area is obtained. Ruggedness is measured in hundreds of meters of elevation difference for grid points 30 arc-seconds apart. Source: *Nunn and Puga (2012)*. (ii) Soil suitability for agriculture The soil suitability component, based on soil carbon density and soil pH, of an index of land suitability for agriculture. The soil suitability data are reported at a half-degree resolution by Ramankutty et al. (2002a) and are aggregated to the country level by Michalopoulos (2012) by averaging across grid cells within a country. For additional details on the soil suitability component of the land suitability index, the interested reader is referred to the definition of the land suitability variable above. Source: *Ashraf and Galor (2013)*. (iii) Mean elevation The mean elevation of a country in km above sea level, calculated using geospatial elevation data reported by the G-ECON project (Nordhaus, 2006) at a 1-degree resolution, which, in turn, is based on similar but more spatially disaggregated data at a 10-minute resolution from New et al. (2002). The measure is thus the average elevation across the grid cells within a country. The interested reader is referred to the G-ECON project website for additional details. Source: *Ashraf and Galor (2013)*. (iv) Mean temperature The intertemporal average monthly temperature of a country in degrees Celsius per month over the 1961–1990 time period, calculated using geospatial average monthly temperature data for this period reported by the G-ECON project (Nordhaus, 2006) at a 1-degree resolution, which, in turn, is based on similar but more spatially disaggregated data at a 10-minute resolution from New et al. (2002). The measure is thus the spatial mean of the intertemporal average monthly temperature across the grid cells within a country. See the G-ECON project website for additional details. Source: *Ashraf and Galor (2013)*. (v) Mean precipitation The intertemporal average monthly precipitation of a country in mm per month over the 1961–1990 time period, calculated using geospatial average monthly precipitation data for this period reported by the G-ECON project (Nordhaus, 2006) at a 1-degree resolution, which, in turn, is based on similar but more spatially disaggregated data at a 10-minute resolution from New et al. (2002). The measure is thus the spatial mean of the intertemporal average monthly precipitation across the grid cells within a country. The interested reader is referred to the G-ECON project web site for additional details. Source: *Ashraf and Galor (2013)*. (vi) Percentage of population living in tropical/subtropical, and temperate zones The percentage of a country's population in 1995 that resided in areas classified as tropical/subtropical, and temperate by the Köppen-Geiger climate classification system. This variable was originally constructed by Gallup et al. (1999) and is part of Harvard University's CID Research Datasets on General Measures of Geography. Source: *Ashraf and Galor (2013)*. (vii) Percentage of land near a waterway The percentage of a country's total land area that is located within 100 km of an ice-free coastline or sea-navigable river. This variable was originally constructed by Gallup et al. (1999) and is part of Harvard University's CID Research Datasets on General Measures of Geography. Source: *Ashraf and Galor (2013)*.
Latitude	The absolute value of the latitude of the country, scaled to take values between 0 and 1. The data is collected from CIA (1996). Source: *La Porta et al. (1999)*.
Log GNI per capita	Logarithm of GNI per capita in current US dollars averaged over the period 1984–2011. Source: *World Bank Development Indicators (WDI), Data retrieved Online in December 2015*.

(*continued on next page*)

Table B1 (*continued*)

Cross country regressions: Variable name	Description and source
Log GNP per capita	Logarithm of GNP per capita in current U.S. dollars averaged over the period 1970–1995. The data is collected from WDI. Source: *La Porta et al. (1999)*.
Log population	Logarithm of population averaged over the period 1984–2011. Source: *World Bank Development Indicators (WDI), Data retrieved Online in December 2015*.
Log population (1960)	Logarithm of population in 1960. This is the variable used in Alesina et al. (2003) for the country size. Our data source might be different. Source: *Penn World Table, Data retrieved Online in December 2015*.
Regional dummy variables	Dummy variable for (1) Sub-Saharan Africa, (2) East Asia Pacific, and (3) Latin America and Caribbean. Source: *World Bank* (http://www.worldbank.org/en/country).
Legal origin dummy variables	Identifies the legal origin of the 212 Company Law or Commercial Code of each country. There are five possible origins: (1) English Common Law; (2) French Commercial Code; (3) German Commercial Code; (4) Scandinavian Commercial Code; and (5) Socialist/Communist Laws. The data is collected using La Porta et al. (1998), American Association of Law Libraries and Flores (1989) and CIA (1996). Source: *La Porta et al. (1999)*
Religion dummy variables	Identifies the percentage of the population of each country that belonged to the three most widely spread religions in the world in 1980. For countries of recent formation the data is available for 1990–1995. The numbers are in percent (scale from 0 to 100). The three religions identified here are: (1) Roman Catholic; (2) Protestant and (3) Muslim. The residual is called ”other religions”. The data is collected using Barrett (1982), Worldmark (1995), Statistical Abstract of the World (1995), UN (1995), CIA (1996). Source: *La Porta et al. (1999)*.

Table B2
Description of the data for within-country analysis (Italy).

Within country (Italy) regressions: Variable name	Description and source
Consanguinity	The data reported by Cavalli-Sforza et al. (2004) on consanguineous marriages, for 5-year periods from 1910 to 1964, comes from the Vatican’s Secret Archives in which requests for dispensations from the consanguinity impediment, sent by the Bishops to the Sacred Congregation of the Sacraments in Rome, were recorded with information of the name of the diocese of marriage, the date and the degree of relationship between the spouses. They grouped 280 Italian dioceses into the provinces present in 1961, for which the number of total marriages was available from year to year. They report four major degrees of consanguinity; uncle-niece/aunt-nephew, first cousins, first cousins once-removed, second cousins. However, the bishops of the islands (Sardinia and Sicily) were granted the privilege to accord the dispensation for consanguineous unions beyond degree III, i.e. first cousins once-removed and second cousins were not recorded in the Vatican Archives, and therefore are not reported for Sicily and are obtained from another source for Sardinia. Thus, we have only considered uncle-niece/aunt-nephew and first cousin unions to compute consanguinity rates for Italian provinces. We have also chosen 5-year periods from 1945 to 1964 for which the data is available in all reported provinces. We computed the consanguinity rate for each province as the average of consanguinity percentages of four 5-year periods (1945–1949, 1950–1954, 1955–1959, 1960–1964) weighted by the number of marriages. For newly created provinces, we used the consanguinity rate of the province they belonged to in year 1961. This gave us consanguinity rates of 108 Italian provinces. Source: *Cavalli-Sforza et al. (2004)*.
Corruption	Province-level number of associative crimes reported by the police to the court per 100,000 inhabitants averaged over 2000–2013. Associative crimes include *criminal association* (article 416: when three or more persons associate together in order to commit more than one crime) and *Mafia-type association* (article 416-bis: participating in a Mafia-type unlawful association including three or more persons). The data is available for 103 Italian provinces. Source: *ISTAT*.
Share of agriculture	Province-level share of agriculture in total value-added averaged over 2000–2013. Source: *ISTAT*.
Log value added per capita	Logarithm of province-level total value added (at current prices millions Euros) averaged over 2000–2013. Source: *ISTAT*.
Log population	Logarithm of province-level population averaged over 2000–2013. Source: *ISTAT*.
Civic involvement	An integer index ranging from 1 to 9 which combines five variables observed in the late 19th century and early 20th century; (i) Membership in mutual aid societies (1873–1904); (ii) Membership in cooperatives (1889–1915); (iii) Strength of the mass parties (1919–1921); (iv) Turnout in the few relatively open elections (1919–1921) before Fascism brought authoritarian rule to Italy; (v) The longevity of local associations founded before 1860. Source: *Putnam et al. (1994)*.
Dominations	A categorical variable that identifies, for each province, the administration that presided during the period of the Spanish domination in Italy, 1560–1659; Spanish, Papal, Austrian, Venetian, Sabaudian and Independent. Source: *Di Liberto and Sideri (2015)*.
Family types	A categorical variable that identifies, for each province, Todd (1990)’s classification of family types which is argued to be very similar to what the geography of family types would have been in the Middle Ages. The three family types common in Italy were the following; (i) incomplete stem family (characterized by an extended family with several generations living under one roof and the inheritance of the house and the land by one son – generally, the eldest – who stays at home); (ii) Communitarian family (characterized by an extended family in which all the sons can get married and bring their wives to the family home and equal division of inheritance among children); (iii) Egalitarian nuclear (characterized with total emancipation of children in adulthood to form independent families and equal division of inheritance among children). Source: Duranton et al. (2009).

(*continued on next page*)

Table B2 (*continued*)

Within country (Italy) regressions:

Variable name	Description and source
Voluntary organizations	Region-level number of voluntary organizations established before 1965 per 100,000 inhabitants at year 2001. The variable takes the same value for all provinces within a region. Source: *ISTAT*.
Family ties	Region-level fraction of youth aged 18–34 living with at least one parent averaged over 2002–2009. The variable takes the same value for all provinces within a region. Source: *ISTAT*.
Alternative corruption metrics	(i) Corruption crimes (region level, N=20) Region-level number of corruption crimes convicted of felony by final judgment per 100,000 inhabitants averaged over 2000–2011. Corruption crimes include the following types of crimes defined under the title *offenses against public administration*: crimes of peculation, malversation, bribery, and corruption. Source: *ISTAT*. (ii) Infrastructure (province level, N=92, and region level, N=20) The difference between a measure of the value of existing physical quantities of public infrastructure and the cumulative price government has paid for public capital stocks. Where the difference is larger between the monies spent and the existing physical infrastructure, more money is being siphoned off to mismanagement, fraud, bribes, kickbacks, and embezzlement; that is, corruption is greater. The measure is created for Italy's 92 provinces and 20 regions as of the mid-1990s, controlling at the regional level for possible differences in the costs of public construction. Inspecting the data, the province-level ratios reported in the appendix appear to be the inverse of the ratios reported in the text for regions, so we have taken the inverse to align the interpretation of the province and region-level measures. Source: *Golden and Picci (2005)*.
South and the islands	A dummy variable for provinces in the following regions: Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, and Sardegna.
Suitability for agriculture	The average of suitability for agriculture in the geographical area of each province. Suitability for agriculture represents the fraction of each grid cell that is suitable to be used for agriculture. It is based on the temperature and soil conditions of each grid cell. The data is construed based on the global map (0.5 by 0.5 ° cells) obtained from Suitability for Agriculture, Atlas of the Biosphere. Source: Ramankutty et al. (2002b).
Distance to coast (in kilometers)	Distance of the geographic center of each province from the coast is constructed based on a coastline physical vector map in 1:10m resolution. Source: Natural Earth.
Ruggedness (in 100 m)	The average of ruggedness in the geographical area of each province. Ruggedness measures the elevation distance of each grid cell and its neighbours. The data is constructed based on the global map (30 by 30 arc-second cells) obtained from Grid-cell-level Data on Terrain Ruggedness. Source: Nunn and Puga (2012).
Elevation (in 100 m)	The average of elevation in the geographical area of each province. Elevation is constructed based on the global map (30 by 30 arc-second cells) obtained from Global 30 Arc-Second Elevation data set. Source: GTOPO30 data set.

Appendix C. Additional cross-country analysis

Table C1
Regression analysis of the relationship between consanguinity and corruption. Higher values of the dependent variable imply lower corruption. The table reports the full regression table underlying Table 2 in Section 3.1.

VARIABLES	(1)	(2)	(3) Restricted sample	(4) Basic model	(5) and religion	(6) both Income and Latitude	(7) and religion
Consanguinity				−3.874***	−2.438**	−2.852**	−1.743
				(0.999)	(1.143)	(1.242)	(1.416)
Ethnic fractionalization	0.232	0.271	−0.380	−0.140	−0.171	−0.352	−0.339
	(0.480)	(0.505)	(0.614)	(0.447)	(0.416)	(0.489)	(0.477)
Latitude		0.628	−3.566	−4.349**	−1.711	−3.782*	−1.307
		(1.747)	(2.357)	(2.092)	(2.114)	(2.081)	(2.129)
Log GNI per capita						0.466	0.395
						(0.344)	(0.343)
Protestant					0.848		0.645
					(0.946)		(1.024)
Catholic					−0.743**		−0.836**
					(0.349)		(0.359)
Muslim					−1.547***		−1.533***
					(0.437)		(0.434)
Log population	−0.041	−0.039	−0.098	−0.126	−0.301	−0.098	−0.273
	(0.113)	(0.114)	(0.204)	(0.182)	(0.204)	(0.166)	(0.186)
Africa	−0.196	−0.172	0.052	0.446	−0.104	0.539	0.022
	(0.251)	(0.262)	(0.538)	(0.365)	(0.357)	(0.387)	(0.371)
East Asia	−0.082	−0.069	−0.274	−0.643	−0.773*	−0.570	−0.729*
	(0.341)	(0.335)	(0.489)	(0.433)	(0.415)	(0.402)	(0.377)

(*continued on next page*)

Please cite this article as: M. Akbari, D. Bahrami-Rad and E.O. Kimbrough, Kinship, fractionalization and corruption, Journal of Economic Behavior and Organization, https://doi.org/10.1016/j.jebo.2019.07.015

Table C1 (*continued*)

VARIABLES	(1)	(2)	(3) Restricted sample	(4) Basic model	(5) and religion	(6) both Income and Latitude	(7) and religion
Latin America	−0.230	−0.207	−0.185	−0.895**	−1.139***	−0.899**	−1.095***
	(0.261)	(0.265)	(0.430)	(0.369)	(0.327)	(0.365)	(0.316)
Socialist legal origin	−1.362***	−1.373***	−0.895**	−0.989***	−0.907***	−0.846**	−0.794***
	(0.285)	(0.287)	(0.360)	(0.318)	(0.285)	(0.323)	(0.287)
French legal origin	−0.080	−0.083	−0.106	−0.047	0.293	0.037	0.362
	(0.174)	(0.174)	(0.266)	(0.227)	(0.227)	(0.252)	(0.255)
German legal origin	0.442	0.440	−0.019	0.122	−0.072	−0.048	−0.288
	(0.427)	(0.420)	(0.415)	(0.388)	(0.470)	(0.417)	(0.506)
Scandinavian legal origin	1.155***	1.110***	0.825*	0.916*	−0.490	0.822*	−0.484
	(0.368)	(0.369)	(0.478)	(0.458)	(0.846)	(0.452)	(0.841)
Temperature	−0.054***	−0.045	−0.145***	−0.115***	−0.071**	−0.109***	−0.068**
	(0.020)	(0.030)	(0.030)	(0.026)	(0.029)	(0.025)	(0.026)
Precipitation	0.002	0.003	−0.001	−0.003	0.001	−0.001	0.002
	(0.003)	(0.003)	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)
Suitability for agriculture	−0.297	−0.316	−0.272	−0.608	−0.869	−0.108	−0.475
	(0.433)	(0.434)	(0.797)	(0.685)	(0.615)	(0.709)	(0.610)
Tropical climate	−0.194	−0.162	0.047	−0.414	−0.396	−0.188	−0.190
	(0.392)	(0.380)	(0.695)	(0.709)	(0.619)	(0.676)	(0.592)
Temperate climate	0.835**	0.826*	0.464	0.142	−0.101	−0.028	−0.215
	(0.415)	(0.427)	(0.792)	(0.616)	(0.620)	(0.625)	(0.649)
Land % near waterways	0.547*	0.532*	0.423	0.314	−0.162	0.032	−0.390
	(0.314)	(0.307)	(0.507)	(0.371)	(0.367)	(0.447)	(0.428)
Elevation	0.469	0.529	−0.322	−0.301	−0.067	−0.125	0.056
	(0.304)	(0.368)	(0.498)	(0.462)	(0.478)	(0.426)	(0.459)
Ruggedness	−0.150	−0.153	−0.279	−0.211	−0.151	−0.231	−0.162
	(0.147)	(0.149)	(0.214)	(0.177)	(0.176)	(0.176)	(0.180)
Constant	3.876***	3.440*	8.284***	9.296***	9.400***	6.618***	7.228***
	(1.002)	(1.749)	(2.439)	(1.833)	(1.606)	(2.355)	(2.121)
Observations	126	126	65	65	65	65	65
R-squared	0.653	0.654	0.680	0.770	0.827	0.781	0.834

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C2
Model specification Table (13), column (3) from Alesina et al. (2003), including consanguinity.

VARIABLES	(1) Alesina (2003)	(2) Alesina (2003) restricted sample	(3) with Consanguinity	(4) with Consanguinity without EF(2003)
Ethnic fractionalization	−1.029	−1.635	−0.225	
	(0.760)	(0.982)	(0.963)	
Consanguinity			−6.076***	−6.251***
			(2.092)	(1.909)
Log GNP per capita	1.028***	1.273***	0.879***	0.868***
	(0.176)	(0.198)	(0.276)	(0.274)
Log population (1960)	0.609***	0.796***	0.347	0.346
	(0.207)	(0.287)	(0.284)	(0.281)
Africa	1.224**	2.371**	2.344**	2.272***
	(0.576)	(0.988)	(0.924)	(0.820)
East Asia	−0.375	−0.100	−0.983	−1.040
	(0.616)	(1.049)	(1.153)	(1.138)
Latin America	−0.672	0.304	−1.230*	−1.285*
	(0.444)	(0.467)	(0.701)	(0.646)
Socialist legal origin	1.065**	0.942	0.546	0.578
	(0.521)	(0.677)	(0.848)	(0.839)
French legal origin	−0.075	−0.331	−0.181	−0.175
	(0.373)	(0.548)	(0.489)	(0.482)
German legal origin	0.103	−1.008	−0.050	0.034
	(0.638)	(0.686)	(0.609)	(0.513)
Scandinavian legal origin	2.157***	1.482**	1.409***	1.468***
	(0.472)	(0.642)	(0.507)	(0.456)
Constant	−3.978**	−6.617***	−1.137	−1.092
	(1.764)	(2.412)	(3.411)	(3.371)
Observations	120	62	62	62
R-squared	0.564	0.623	0.695	0.695

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C3
Regression analysis of the relationship between consanguinity and corruption, controlling for date of data collection. Higher values of the dependent variable imply lower corruption. N varies due to missing data for some countries.

VARIABLES	(1) Full sample 1922–2013	(2) Data collection date 1950–2013	(3) Data collection date 1960–2013	(4) Data collection date 1970–2013
Consanguinity	−4.815***	−5.174***	−5.435***	−5.681***
	(1.012)	(1.091)	(0.874)	(1.287)
Consanguinity data collection date	−0.004			
	(0.003)			
Ethnic fractionalization	0.049	0.009	0.692	0.615
	(0.525)	(0.502)	(0.549)	(0.924)
Latitude	−3.871*	−3.888*	−4.691**	−12.291
	(2.207)	(2.157)	(2.087)	(7.914)
Log population	−0.091	−0.061	0.048	0.191
	(0.221)	(0.218)	(0.264)	(0.484)
Africa	−0.167	−0.218	−0.560*	
	(0.339)	(0.377)	(0.313)	
East Asia	−0.800	−0.760	−0.429	−1.802*
	(0.519)	(0.586)	(0.457)	(0.788)
Latin America	−0.969**	−0.909*	−0.922*	−1.804***
	(0.415)	(0.457)	(0.490)	(0.469)
Socialist legal origin	−0.899*	−1.135***	−1.261***	−2.099*
	(0.457)	(0.387)	(0.313)	(1.057)
French legal origin	−0.264	−0.255	−0.529*	−0.567
	(0.255)	(0.271)	(0.277)	(0.315)
German legal origin	0.075	0.082		
	(0.466)	(0.481)		
Scandinavian legal origin	0.590	0.570	1.238***	0.022
	(0.524)	(0.638)	(0.419)	(1.968)
Temperature	−0.101***	−0.103***	−0.089**	−0.313**
	(0.031)	(0.029)	(0.035)	(0.126)
Precipitation	−0.005	−0.006	−0.008*	−0.010
	(0.005)	(0.005)	(0.004)	(0.012)
Suitability for agriculture	−1.422*	−1.476*	−1.009	−1.990
	(0.793)	(0.726)	(0.838)	(1.563)
Tropical climate	0.091	0.100	−0.323	−0.729
	(0.695)	(0.694)	(0.917)	(2.454)
Temperate climate	0.311	0.269	0.455	−0.461
	(0.760)	(0.794)	(0.723)	(1.589)
Land % near waterways	0.492	0.472	0.362	0.994
	(0.382)	(0.410)	(0.423)	(0.797)
Elevation	−0.074	−0.183	−0.198	−0.478
	(0.520)	(0.492)	(0.479)	(0.922)
Terrain Ruggedness	−0.103	−0.071	−0.077	−0.190
	(0.195)	(0.206)	(0.143)	(0.373)
Constant	16.275**	8.972***	8.033***	15.079**
	(6.067)	(1.906)	(2.119)	(4.801)
Observations	56	52	42	26
R-squared	0.812	0.789	0.875	0.937

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

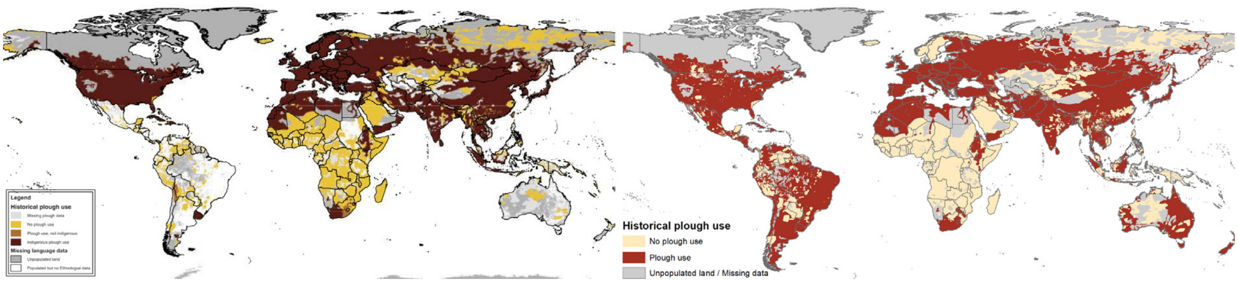

Fig. C1. Traditional plough use across ethnic/linguistic groups shows very high similarity with the map from Alesina et al. (2013) except for the newly added polygons in Australasia, central America, and south America (since we used new version of polygons). Our replicated country-level measure for plough agriculture is highly correlated with the measure from Alesina et al. (2013) (Pearson's $r = 0.72$, p-value < 0.001, $N = 218$) despite using a newer version of the Ethnographic Atlas data and different population data.

Table C4
Regression analysis of the relationship between consanguinity and corruption, using alternative measures of corruption: corruption perception index (1998–2015) ranging 0–100 from Transparency International, control of corruption (1996–2017) ranging −2.5–2.5 from World Bank Group, and Transparency International's historical data (1980–1985) ranging 1–10.

VARIABLES	(1) Corruption perception index	(2) Control of corruption	(3) TI' historical corruption data
Consanguinity	−51.144**	−2.579**	−11.136***
	(21.978)	(1.093)	(3.042)
Ethnic fractionalization	−4.348	−0.086	0.387
	(8.748)	(0.426)	(2.549)
Latitude	−61.804	−2.461	−13.464*
	(42.050)	(1.980)	(7.010)
Log population	−4.201	−0.168	0.203
	(3.006)	(0.157)	(1.079)
Africa	2.970	0.233	−0.192
	(7.541)	(0.395)	(1.301)
East Asia	−0.089	−0.068	−1.777
	(8.081)	(0.408)	(1.928)
Latin America	−3.924	−0.124	−2.626
	(8.221)	(0.397)	(1.887)
Socialist legal origin	−24.242***	−0.959***	−4.268**
	(6.601)	(0.326)	(1.714)
French legal origin	−9.547**	−0.451**	−1.466
	(3.864)	(0.194)	(1.229)
German legal origin	3.717	0.104	−0.413
	(7.386)	(0.370)	(2.338)
Scandinavian legal origin	10.032	0.503	0.057
	(8.770)	(0.426)	(1.985)
Temperature	−1.720***	−0.068**	−0.198*
	(0.579)	(0.028)	(0.106)
Precipitation	−0.084	−0.003	−0.009
	(0.082)	(0.004)	(0.019)
Suitability for agriculture	−18.435	−0.789	−5.942
	(12.679)	(0.684)	(3.794)
Tropical climate	−13.033	−0.753	−2.132
	(14.607)	(0.723)	(4.214)
Temperate climate	4.298	0.287	2.580
	(10.822)	(0.565)	(3.134)
Land % near waterways	17.265**	0.811*	2.287
	(8.306)	(0.404)	(2.599)
Elevation	−8.195	−0.375	2.308
	(6.852)	(0.338)	(2.707)
Ruggedness	−3.257	−0.146	−0.504
	(2.711)	(0.139)	(0.821)
Constant	163.192***	4.872***	15.929**
	(35.986)	(1.746)	(7.099)
Observations	65	65	36
R-squared	0.783	0.758	0.794

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

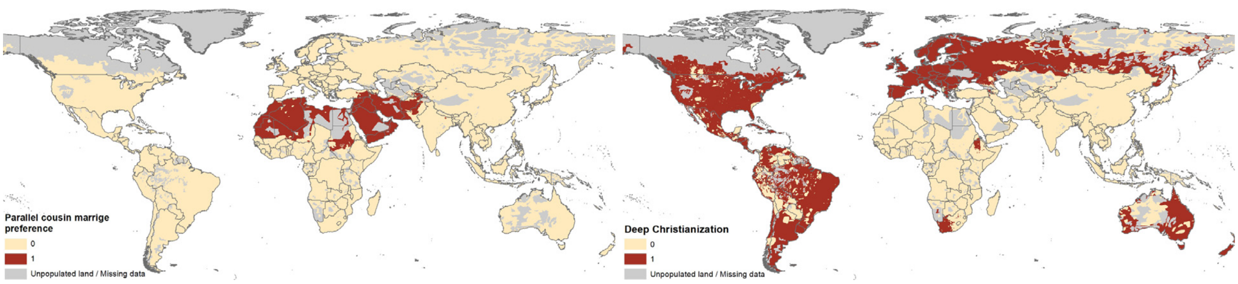

Fig. C2. (Left) Traditional parallel cousin marriage preference. (Right) Deep Christianization across ethnic/linguistic groups.

Table C5
Regression analysis of the relationship between consanguinity and corruption, including additional confounding factors. Higher values of the dependent variable imply lower corruption.

VARIABLES	(1) Basic model	(2) and Genetic diversity	(3) Trust	(4) and Family ties
Consanguinity	−3.874***	−2.834***	−3.211***	−2.777**
	(0.999)	(0.988)	(1.069)	(1.019)
Ethnic fractionalization	−0.140	−0.360	−0.023	−0.010
	(0.447)	(0.395)	(0.429)	(0.415)
Genetic diversity		142.227**	119.934	55.090
		(65.867)	(79.010)	(64.616)
Genetic diversity squared		−108.358**	−90.100	−44.355
		(49.716)	(59.214)	(48.230)
General trust			−0.248	−0.351
			(1.151)	(0.862)
Family ties				−1.707**
				(0.654)
Log population	−0.126	−0.157	−0.157	−0.284
	(0.182)	(0.195)	(0.235)	(0.199)
Latitude	−4.349**	−4.249**	−3.781	−5.556*
	(2.092)	(2.006)	(2.405)	(2.761)
Africa	0.446	0.751*	0.647	0.140
	(0.365)	(0.373)	(0.537)	(0.592)
East Asia	−0.643	−0.813*	−0.294	−0.521
	(0.433)	(0.448)	(0.526)	(0.531)
Latin America	−0.895**	−0.830	−0.445	−1.024
	(0.369)	(0.520)	(0.850)	(0.829)
Socialist legal origin	−0.989***	−0.871**	−1.167**	−1.252**
	(0.318)	(0.332)	(0.447)	(0.564)
French legal origin	−0.047	0.021	−0.169	−0.078
	(0.227)	(0.229)	(0.343)	(0.437)
German legal origin	0.122	−0.334	0.030	−0.612
	(0.388)	(0.475)	(0.518)	(0.639)
Scandinavian legal origin	0.916*	0.948*	1.039	0.134
	(0.458)	(0.485)	(0.773)	(0.671)
Temperature	−0.115***	−0.106***	−0.090**	−0.101***
	(0.026)	(0.026)	(0.033)	(0.035)
Precipitation	−0.003	−0.004	−0.006	−0.013**
	(0.005)	(0.005)	(0.005)	(0.005)
Suitability for agriculture	−0.608	−0.880	−0.720	−1.412*
	(0.685)	(0.708)	(0.841)	(0.727)
Tropical climate	−0.414	−0.309	−0.402	0.940
	(0.709)	(0.654)	(0.783)	(0.691)
Temperate climate	0.142	0.602	0.846	1.037*
	(0.616)	(0.643)	(0.601)	(0.520)
Land % near waterways	0.314	0.414	0.266	0.303
	(0.371)	(0.382)	(0.491)	(0.495)
Elevation	−0.301	−0.026	0.124	−0.434
	(0.462)	(0.475)	(0.516)	(0.584)
Ruggedness	−0.211	−0.246	−0.338**	0.113
	(0.177)	(0.173)	(0.164)	(0.231)
Constant	9.296***	−36.974*	−30.484	−5.079
	(1.833)	(21.428)	(25.659)	(21.380)
Observations	65	65	54	44
R-squared	0.770	0.795	0.842	0.929
Geographical variables	yes	yes	yes	yes

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C6
Regression analysis of the relationship between lineage endogamy and corruption. Higher values of the dependent variable imply lower corruption. The table reports the full regression table underlying Table 4 in Section 3.1.

VARIABLES	(1) Basic model	(2) both Income and Latitude	(3) and Religion
Lineage endogamy	−1.022**	−0.898**	−0.702*
	(0.440)	(0.356)	(0.369)
Ethnic fractionalization	0.202	0.209	0.225
	(0.479)	(0.424)	(0.412)
Latitude	0.664	0.617	0.747
	(1.716)	(1.615)	(1.570)
Log GNI per capita		0.669***	0.547**
		(0.199)	(0.215)
Protestant			0.685
			(0.468)
Catholic			0.045
			(0.260)
Muslim			−0.375
			(0.379)
Log population	−0.110	0.010	0.028
	(0.111)	(0.100)	(0.107)
Africa	−0.598*	−0.072	−0.215
	(0.348)	(0.283)	(0.286)
East Asia	−0.170	−0.077	−0.177
	(0.323)	(0.309)	(0.318)
Latin America	−0.499*	−0.446*	−0.579**
	(0.290)	(0.248)	(0.269)
Socialist legal origin	−1.478***	−1.123***	−1.025***
	(0.286)	(0.301)	(0.294)
French legal origin	0.001	0.045	0.108
	(0.172)	(0.168)	(0.176)
German legal origin	0.371	0.075	0.098
	(0.374)	(0.341)	(0.317)
Scandinavian legal origin	0.937**	1.009***	0.568
	(0.374)	(0.336)	(0.435)
Temperature	−0.031	−0.020	−0.013
	(0.031)	(0.030)	(0.030)
Precipitation	−0.000	0.001	0.001
	(0.003)	(0.003)	(0.003)
Suitability for agriculture	−0.809*	−0.271	−0.413
	(0.482)	(0.441)	(0.433)
Tropical climate	−0.141	−0.113	−0.225
	(0.376)	(0.338)	(0.397)
Temperate climate	0.663	0.363	0.271
	(0.401)	(0.376)	(0.418)
Land % near waterways	0.526*	0.340	0.392
	(0.295)	(0.283)	(0.289)
Elevation	0.476	0.460	0.407
	(0.359)	(0.319)	(0.330)
Ruggedness	−0.079	−0.064	−0.064
	(0.131)	(0.117)	(0.111)
Constant	4.531***	0.739	1.024
	(1.724)	(1.606)	(1.686)
Observations	126	125	124
R-squared	0.680	0.716	0.733

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Please cite this article as: M. Akbari, D. Bahrami-Rad and E.O. Kimbrough, Kinship, fractionalization and corruption, Journal of Economic Behavior and Organization, https://doi.org/10.1016/j.jebo.2019.07.015

Table C7
Regression analysis of the relationship between deep Christianization and corruption. Higher values of the dependent variable imply lower corruption. The table includes the full regression table underlying Table 5 in Section 3.1.

	Reduced form			First stage			IV estimation		
VARIABLES	(1) Basic model	(2) both Income and Latitude	(3) and Religion	(4) Basic model	(5) both Income and Latitude	(6) and Religion	(7) Basic model	(8) both Income and Latitude	(9) and Religion
Deep Christianization	0.986***	0.741**	0.618*	−0.248***	−0.147***	−0.082			
	(0.300)	(0.289)	(0.343)	(0.052)	(0.053)	(0.068)			
Consanguinity							−6.666***	−8.926**	−14.395
							(1.315)	(3.723)	(13.344)
Ethnic fractionalization	0.049	0.085	0.087	0.074	0.105	0.101	−0.008	0.331	0.913
	(0.444)	(0.412)	(0.414)	(0.066)	(0.063)	(0.062)	(0.489)	(0.858)	(1.774)
Latitude	−0.413	−0.346	−0.121	−0.164	−0.236	−0.330	−5.252**	−6.018*	−6.703
Log GNI per capita		0.494**	0.458**		−0.105**	−0.097*		−0.577	−1.122
		(0.201)	(0.212)		(0.052)	(0.048)		(0.821)	(1.821)
Protestant			0.634			−0.178			−1.793
			(0.487)			(0.158)			(3.732)
Catholic			−0.152			0.002			−1.323*
			(0.301)			(0.048)			(0.741)
Muslim			−0.310			0.064			−0.460
			(0.362)			(0.086)			(1.935)
Log population	−0.138	−0.039	−0.014	0.023	0.008	0.014	−0.165	−0.198	−0.246
	(0.106)	(0.103)	(0.109)	(0.023)	(0.020)	(0.020)	(0.205)	(0.240)	(0.345)
	(1.528)	(1.520)	(1.506)	(0.329)	(0.303)	(0.328)	(2.448)	(3.096)	(7.128)
Africa	−0.192	0.157	0.030	0.114	0.066	0.104	0.709	0.704	1.208
	(0.256)	(0.252)	(0.264)	(0.096)	(0.097)	(0.070)	(0.639)	(0.793)	(1.294)
East Asia	0.292	0.257	0.109	−0.189***	−0.145***	−0.113*	−0.907*	−1.092*	−1.704
	(0.306)	(0.311)	(0.336)	(0.055)	(0.050)	(0.060)	(0.493)	(0.648)	(1.196)
Latin America	−0.726**	−0.595**	−0.594**	−0.007	−0.035	−0.035	−1.435***	−1.602***	−1.688*
	(0.279)	(0.246)	(0.265)	(0.048)	(0.042)	(0.048)	(0.379)	(0.508)	(0.859)
Socialist legal origin	−1.520***	−1.244***	−1.131***	0.065	0.005	−0.021	−1.096***	−1.280***	−1.526
	(0.261)	(0.292)	(0.294)	(0.042)	(0.039)	(0.042)	(0.349)	(0.457)	(0.925)
French legal origin	−0.099	−0.052	0.062	0.010	−0.010	−0.026	−0.009	−0.095	−0.021
	(0.166)	(0.171)	(0.178)	(0.036)	(0.030)	(0.037)	(0.280)	(0.327)	(0.822)
German legal origin	0.370	0.133	0.101	−0.200***	−0.068	−0.031	0.146	0.424	−0.062
	(0.313)	(0.300)	(0.292)	(0.066)	(0.071)	(0.081)	(0.423)	(0.603)	(1.024)
Scandinavian legal origin	0.880**	0.967***	0.479	0.057	0.061	0.185	0.955*	1.106	1.622
	(0.353)	(0.334)	(0.441)	(0.058)	(0.056)	(0.146)	(0.544)	(0.710)	(3.458)

(*continued on next page*)

Table C7 (*continued*)

VARIABLES	Reduced form			First stage			IV estimation		
	(1) Basic model	(2) both Income and Latitude	(3) and Religion	(4) Basic model	(5) both Income and Latitude	(6) and Religion	(7) Basic model	(8) both Income and Latitude	(9) and Religion
Temperature	−0.036	−0.030	−0.022	0.003	0.001	−0.001	−0.095***	−0.095**	−0.085
	(0.027)	(0.029)	(0.030)	(0.005)	(0.004)	(0.006)	(0.035)	(0.043)	(0.089)
Precipitation	−0.000	0.001	0.001	0.000	−0.000	−0.001	−0.004	−0.007	−0.009
	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.006)	(0.007)	(0.016)
Suitability for agriculture	−0.620	−0.187	−0.297	−0.120*	−0.196**	−0.167**	−0.892	−1.579	−2.807
	(0.389)	(0.404)	(0.405)	(0.065)	(0.076)	(0.074)	(0.751)	(1.320)	(2.812)
Tropical climate	0.126	0.096	0.032	−0.140	−0.163*	−0.151	−0.637	−1.080	−1.641
	(0.353)	(0.326)	(0.393)	(0.105)	(0.093)	(0.114)	(0.986)	(1.471)	(2.290)
Temperate climate	0.779**	0.594*	0.571	0.005	0.014	−0.009	0.075	0.136	0.176
	(0.376)	(0.350)	(0.390)	(0.082)	(0.079)	(0.088)	(0.665)	(0.783)	(1.265)
Land % near waterways	0.388	0.292	0.359	−0.015	0.050	0.071	0.229	0.554	0.740
	(0.284)	(0.281)	(0.288)	(0.048)	(0.049)	(0.054)	(0.358)	(0.569)	(1.481)
Elevation	0.562*	0.527*	0.518	−0.095	−0.095	−0.093	−0.289	−0.501	−0.870
	(0.334)	(0.310)	(0.332)	(0.060)	(0.064)	(0.063)	(0.500)	(0.735)	(1.512)
Ruggedness	−0.051	−0.044	−0.039	0.009	0.009	0.004	−0.112	−0.090	0.021
	(0.117)	(0.112)	(0.108)	(0.025)	(0.025)	(0.026)	(0.213)	(0.250)	(0.379)
Constant	4.388***	1.589	1.493	0.208	0.752**	0.721**	10.276***	13.746**	18.450
	(1.563)	(1.697)	(1.702)	(0.318)	(0.369)	(0.324)	(2.082)	(5.200)	(12.375)
Observations	125	124	123	64	64	64	64	64	64
R-squared	0.720	0.736	0.750	0.865	0.884	0.893	0.719	0.633	0.304

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Please cite this article as: M. Akbari, D. Bahrami-Rad and E.O. Kimbrough, Kinship, fractionalization and corruption, Journal of Economic Behavior and Organization, https://doi.org/10.1016/j.jebo.2019.07.015

Appendix D. Additional within-country analyses (Italy)

Table D1
Regression analysis of the relationship between consanguinity and associative crimes in Italy: controlling for confounding factors.

VARIABLES	(1) Basic model	(2) and family types	(3) and family ties	(4) and civic involvement (1860–1921)	(5) and voluntary organizations (before 1965)	(6) and dominations (Middle ages)	(7) and South	(8) and latitude
Consanguinity	0.582**	0.575**	0.601***	0.508**	0.503**	0.536**	0.568**	0.508
	(0.211)	(0.213)	(0.201)	(0.198)	(0.200)	(0.228)	(0.251)	(0.339)
Family ties		0.023	0.026	−0.005	−0.007	0.003	0.018	0.022
		(0.040)	(0.038)	(0.058)	(0.059)	(0.069)	(0.068)	(0.065)
Communitarian family			0.655**	0.673**	0.709**	0.512	0.762	0.712
			(0.279)	(0.263)	(0.301)	(0.492)	(0.617)	(0.632)
Egalitarian nuclear family			0.522*	0.333	0.365	0.329	0.436	0.409
			(0.297)	(0.349)	(0.376)	(0.414)	(0.439)	(0.444)
Civic involvement				−0.111	−0.119	−0.155	−0.212	−0.210
				(0.102)	(0.107)	(0.121)	(0.154)	(0.155)
Voluntary organizations					0.004	0.005	0.010	0.009
					(0.007)	(0.010)	(0.012)	(0.011)
Papal						0.476	0.097	−0.163
						(0.409)	(0.462)	(0.566)
Austrian						−0.217	−0.365	−0.406
						(0.414)	(0.324)	(0.346)
Venetian						0.278	0.209	0.157
						(0.215)	(0.158)	(0.189)
Sabaudian						0.134	0.005	−0.005
						(0.430)	(0.397)	(0.395)
Independent						0.496	0.281	0.192
						(0.424)	(0.430)	(0.471)
South and Islands							−0.698	−1.055
							(0.910)	(0.913)
Latitude								−0.132
								(0.253)
Share of agriculture	1.589	1.040	1.509	1.632	1.443	1.072	1.021	0.866
	(3.047)	(3.023)	(2.918)	(2.887)	(2.992)	(3.117)	(3.158)	(3.242)
Log population	−0.018	−0.019	0.007	−0.029	−0.036	−0.048	−0.043	−0.048
	(0.092)	(0.090)	(0.095)	(0.100)	(0.105)	(0.108)	(0.111)	(0.112)
Mean Temperature	−0.044	−0.057	−0.096	−0.079	−0.080	−0.083	−0.083	−0.113
	(0.062)	(0.058)	(0.067)	(0.065)	(0.066)	(0.074)	(0.076)	(0.090)
Mean precipitation	−0.902	−0.436	0.558	−0.178	−0.040	0.127	0.649	−0.832
	(3.573)	(3.680)	(3.384)	(3.635)	(3.699)	(3.774)	(3.327)	(4.371)
Suitability for agriculture	0.097	0.085	0.250	0.151	0.202	0.147	0.113	0.224
	(0.405)	(0.380)	(0.399)	(0.463)	(0.525)	(0.676)	(0.670)	(0.670)
Distance to coast	−0.006***	−0.005**	−0.006**	−0.004	−0.004	−0.002	−0.003	−0.003
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Elevation	0.094	0.075	0.065	0.026	0.021	0.027	0.029	0.026
	(0.074)	(0.069)	(0.064)	(0.063)	(0.067)	(0.074)	(0.073)	(0.070)
Ruggedness	−0.205	−0.183	−0.158	−0.077	−0.071	−0.109	−0.107	−0.098
	(0.140)	(0.141)	(0.141)	(0.138)	(0.143)	(0.153)	(0.154)	(0.151)
Constant	1.668	0.355	−0.817	2.381	2.538	2.144	1.758	7.506
	(1.243)	(2.677)	(2.489)	(4.733)	(4.849)	(5.787)	(5.722)	(14.650)
Observations	101	101	101	101	101	101	101	101
R-squared	0.475	0.480	0.504	0.516	0.517	0.530	0.533	0.535

Robust standard errors, clustered at region level, in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table D2
Robustness check for date of data collection, Italy.

VARIABLES	(1) Consanguinity sample of basic model (1945–1964)	(2) Consanguinity sample restricted to 1950–1964	(3) Consanguinity sample restricted to 1955–1964	(4) Consanguinity sample restricted to 1960–1964
Consanguinity	0.582**	6.225**	6.456**	6.675**
	(0.211)	(2.353)	(2.616)	(2.878)
Share of agriculture	1.589	1.370	1.254	0.884
	(3.047)	(3.034)	(3.011)	(3.034)
Log population	−0.018	−0.019	−0.015	−0.012
	(0.092)	(0.089)	(0.084)	(0.083)
Mean Temperature	−0.044	−0.042	−0.041	−0.034
	(0.062)	(0.061)	(0.061)	(0.060)
Mean precipitation	−0.902	−1.171	−1.360	−1.705
	(3.573)	(3.818)	(4.209)	(4.697)
Suitability for agriculture	0.097	0.063	0.074	0.074
	(0.405)	(0.398)	(0.391)	(0.387)
Distance to coast	−0.006***	−0.006***	−0.006***	−0.006***
	(0.002)	(0.002)	(0.002)	(0.002)
Elevation	0.094	0.094	0.095	0.102
	(0.074)	(0.073)	(0.073)	(0.072)
Ruggedness	−0.205	−0.204	−0.202	−0.208
	(0.140)	(0.137)	(0.135)	(0.134)
Constant	1.668	1.745	1.714	1.714
	(1.243)	(1.202)	(1.149)	(1.127)
Observations	101	101	101	101
R-squared	0.475	0.480	0.485	0.483

Robust standard errors, clutered at region level, in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table D3
Replication of Table 6 using the corruption measure from Golden and Picci (2005).

VARIABLES	(1) Basic model	(2) and Consanguinity	(3) both Income and Share of agriculture
Consanguinity		−0.144*	−0.153*
		(0.077)	(0.078)
Share of agriculture	0.475	1.608	−0.656
	(1.308)	(1.304)	(1.929)
Log value added per capita			−0.107*
			(0.051)
Log population	−0.163*	−0.115	−0.198*
	(0.093)	(0.086)	(0.100)
Mean Temperature	−0.053	−0.032	−0.030
	(0.031)	(0.031)	(0.030)
Mean precipitation	−1.614	1.195	1.564
	(3.189)	(3.136)	(3.178)
Suitability for agriculture	0.250	0.274	0.283
	(0.478)	(0.461)	(0.438)
Distance to coast	0.004	0.004	0.004
	(0.003)	(0.003)	(0.002)
Elevation	−0.001	−0.011	−0.015
	(0.054)	(0.056)	(0.056)
Ruggedness	−0.084	−0.032	−0.038
	(0.117)	(0.128)	(0.127)
Constant	3.162**	2.462*	3.219**
	(1.387)	(1.323)	(1.352)
Observations	90	90	90
R-squared	0.328	0.352	0.366

Robust standard errors, clutered at region level, in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table D4
Replication of Table 7 using the corruption measure from Golden and Picci (2005).

	(1)	(2)	(3)	(4)	(5)	(6)
	Reduced form		First stage		IV estimation	
VARIABLES	Basic model	both Income and Share of agriculture	Basic model	both Income and Share of agriculture	Basic model	both Income and Share of agriculture
Exposure to the Church	0.096***	0.094**	−0.168***	−0.171***		
	(0.034)	(0.033)	(0.047)	(0.049)		
Consanguinity					−0.572**	−0.548**
					(0.229)	(0.213)
Share of agriculture	2.419*	0.625	4.448	2.156	4.966**	1.805
	(1.393)	(2.151)	(3.246)	(3.917)	(1.933)	(2.142)
Log value added per capita		−0.080		−0.102		−0.136***
		(0.052)		(0.074)		(0.050)
Log population	−0.122	−0.187**	0.259***	0.176*	0.026	−0.091
	(0.078)	(0.088)	(0.081)	(0.101)	(0.085)	(0.084)
Mean Temperature	−0.029	−0.029	0.102***	0.102***	0.030	0.026
	(0.031)	(0.030)	(0.031)	(0.031)	(0.043)	(0.040)
Mean precipitation	1.187	1.261	14.551***	14.646***	9.517**	9.280**
	(2.909)	(2.907)	(3.269)	(3.308)	(4.111)	(3.889)
Suitability for agriculture	−0.126	−0.111	0.828*	0.847*	0.348	0.353
	(0.437)	(0.429)	(0.483)	(0.468)	(0.410)	(0.377)
Distance to coast	0.003	0.004*	0.000	0.001	0.004	0.004*
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Elevation	0.005	0.002	−0.079	−0.083	−0.041	−0.044
	(0.057)	(0.058)	(0.052)	(0.051)	(0.061)	(0.060)
Ruggedness	−0.089	−0.096	0.370***	0.361***	0.123	0.102
	(0.115)	(0.116)	(0.119)	(0.121)	(0.146)	(0.140)
Constant	1.937*	2.566**	−2.708*	−1.904	0.387	1.524
	(1.097)	(1.152)	(1.410)	(1.625)	(1.250)	(1.167)
Observations	90	90	90	90	90	90
R-squared	0.392	0.400	0.829	0.832	0.149	0.195

Robust standard errors, clustered at region level, in parentheses.*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

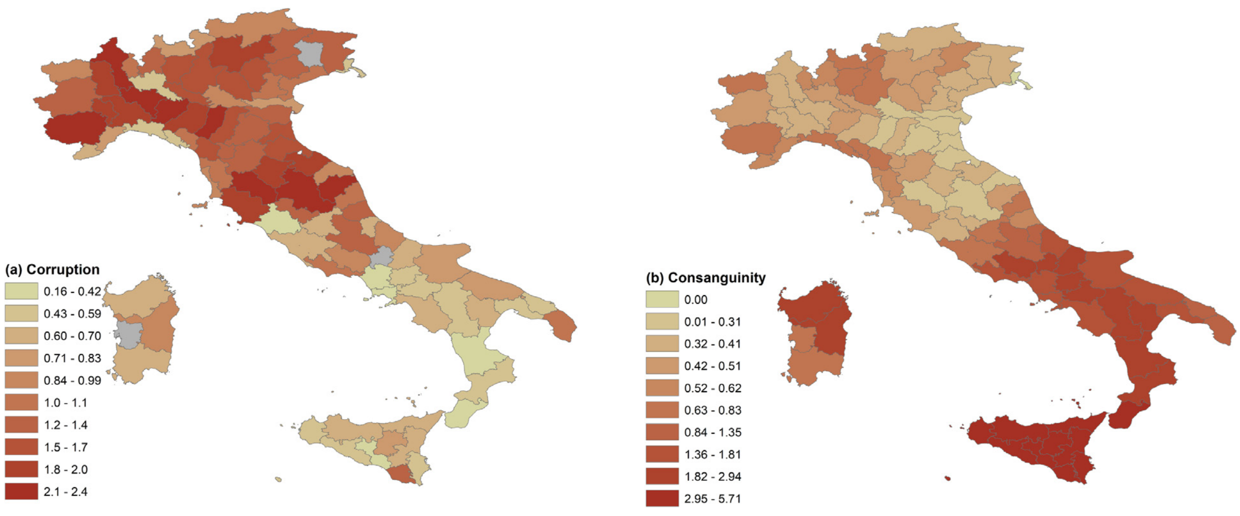

Fig. D1. Alternative measure of corruption from Golden and Picci (2005). Darker colored regions are less corrupt. Grey colored areas indicate missing data. See Appendix B for details on the data. Here a higher score is associated with lower corruption (in contrast to our baseline measure which uses the number of crimes). This alternative measure of corruption is highly correlated with consanguinity (Pearson's $r=-0.52$, p-value <0.001, N = 90). Moreover, Tables D.3 and D.4 reports replications of our province-level regressions from Tables 6 and 7 in the body of the paper using this alternative measure of corruption. Consanguinity remains a robust predictor of corruption in all specifications. However, as Golden and Picci (2005) note, their province-level ratios are not cost-adjusted due to lack of data (the region-level ratios are cost-adjusted), and this is why we focus our main analysis on associative crime. Finally, note also that our measures of share of agriculture in GDP and population come from the period 2000–2013 (and hence from a period after Golden and Picci's data on corruption). We do not have access to earlier province-level statistics on GDP or population, so the reliability of these specifications depends on the assumption that provincial GDP (and share of agriculture therein) and population in 2000–2013 is highly correlated with the same in 1954–1997, which seems plausible.

Appendix E. Additional analyses with European regions

Table E1
Regression analyses of the relationship between consanguinity and corruption in European regions (Italy, France, Spain): controlling for confounding factors.

VARIABLES	(1) Basic model	(2) Trust	(3) Religion	(4) Fractionalization	(5) Civicness
Consanguinity	−5.198***	−3.128***	−2.645*	−2.798**	−2.419*
	(1.536)	(1.160)	(1.364)	(1.295)	(1.426)
Generalized trust		0.972***	0.941***	0.880***	0.860***
		(0.124)	(0.118)	(0.112)	(0.125)
Protestant (%)			−1.111***	−1.000***	−1.032***
			(0.346)	(0.325)	(0.335)
Party fractionalization				−20.365**	−18.965**
				(8.938)	(8.981)
Voter turnout					0.072
					(0.174)
Log population	−0.957	−1.968***	−1.739**	−1.457**	−1.400*
	(0.836)	(0.594)	(0.696)	(0.714)	(0.823)
Latitude	1.125	0.278	−0.118	0.288	0.193
	(0.722)	(0.474)	(0.559)	(0.630)	(0.633)
Mean temperature	−0.444	−0.003	−0.301	0.012	−0.101
	(0.600)	(0.434)	(0.465)	(0.501)	(0.561)
Mean precipitation	0.078	0.028	0.004	0.023	0.017
	(0.047)	(0.029)	(0.039)	(0.042)	(0.039)
Suitability for agriculture	−0.003	−0.005**	−0.006***	−0.005**	−0.005**
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Distance to coast	−0.000	0.000	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Elevation	−0.003	−0.006	−0.011*	−0.009*	−0.011*
	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
Ruggedness	0.000	0.000	0.000*	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	32.093	40.303	58.577*	47.253	46.085
	(36.694)	(25.119)	(29.535)	(31.411)	(32.405)
Country FEs	yes	yes	yes	yes	yes
Observations	57	56	54	54	54
R-squared	0.883	0.951	0.955	0.960	0.960

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table E2
Regression analyses of the relationship between consanguinity and corruption in European regions (Italy, France, Spain): IV estimations.

	IV estimation		
VARIABLES	(1)	(2)	(3)
Consanguinity	−8.783**	−9.524***	−7.864
	(3.935)	(3.244)	(5.470)
Log population	−3.330***	−2.210*	−1.147
	(1.137)	(1.236)	(1.073)
Latitude		–-	1.274
			(0.861)
Mean temperature		–	−0.127
			(0.917)
Mean precipitation		–	0.129
			(0.124)
Suitability for agriculture		−0.006	−0.002
		(0.004)	(0.003)

(*continued on next page*)

Please cite this article as: M. Akbari, D. Bahrami-Rad and E.O. Kimbrough, Kinship, fractionalization and corruption, Journal of Economic Behavior and Organization, https://doi.org/10.1016/j.jebo.2019.07.015

Table E2 (*continued*)

	IV estimation		
VARIABLES	(1)	(2)	(3)
Distance to coast		0.000	−0.000
		(0.000)	(0.000)
Elevation		−0.009	−0.003
		(0.007)	(0.007)
Ruggedness		0.000*	0.000
		(0.000)	(0.000)
Constant	118.022***	110.844***	27.631
	(20.869)	(25.623)	(42.096)
Country FEs	yes	yes	yes
Observations	57	57	57
R-squared	0.795	0.828	0.875

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2019.07.015.

References

Akrami, S.M., Osati, Z., 2007. Is consanguineous marriage religiously encouraged? Islamic and Iranian considerations. J. Biosoc. Sci. 39 (02), 313–316.
Alesina, A., Algan, Y., Cahuc, P., Giuliano, P., 2015. Family values and the regulation of labor. J. Eur. Econ. Assoc 13 (4), 599–630.
Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., Wacziarg, R., 2003. Fractionalization. J. Econ. Growth 8 (2), 155–194.
Alesina, A., Giuliano, P., 2010. The power of the family. J. Econ. Growth 15 (2), 93–125.
Alesina, A., Giuliano, P., 2011. Family ties and political participation. J. Eur. Econ. Assoc. 9 (5), 817–839.
Alesina, A., Giuliano, P., 2014. Family ties. In: Aghion, P., Durlauf, S.N. (Eds.), Handbook of Economic Growth, Vol. 2a, pp. 177–215. North Holland: Netherlands.
Alesina, A., Giuliano, P., Nunn, N., 2013. On the origins of gender roles: women and the plough. Q. J. Econ. qjt005.
Alesina, A., La Ferrara, E., 2002. Who trusts others? J. Public Econ. 85 (2), 207–234.
Ashraf, Q., Galor, O., 2013. The 'Out of Africa' hypothesis, human genetic diversity, and comparative economic development. Am. Econ. Rev. 103 (1), 1.
Augustine, S., 1998. The City of God Against the Pagans. Cambridge University Press. (trans: R. W. Dyson)
Bahrami-Rad, D., 2018. Keeping it in the family: female inheritance, inmarriage, and the status of women. Marriage, and the Status of Women.
Bailey, D.H., Hill, K.R., Walker, R.S., 2014. Fitness consequences of spousal relatedness in 46 small-scale societies. Biol. Lett. 10 (5), 20140160.
Banfield, E.C., 1958. The moral basis of a backward society. N. Y. Glencol. 3, 4.
Barrett, D.B., 1982. World Christian Encyclopedia: A Comparative Study of Churches and Religions in the Modern World, AD 1900–2000. Oxford University Press, USA.
Bigoni, M., Bortolotti, S., Casari, M., Gambetta, D., Pancotto, F., 2016. Amoral familism, social capital, or trust? The behavioural foundations of the Italian north–south divide. Econ. J. 126 (594), 1318–1341.
Bittles, A., Black, M. H., 2015. Global Patterns & Tables of Consanguinity.
Bittles, A.H., 2012. Consanguinity in Context, vol. 63. Cambridge University Press.
Bondarenko, D., Kazankov, A., Khaltourina, D., Korotayev, A., 2005. Ethnographic atlas XXXi: peoples of easternmost europe. Ethnology 261–289.
Bruk, S., 1964. In: Apenchenko, V. (Ed.), Atlas narodov. Moscow: Miklukho-Maklai Ethnological Institute at the Department of Geodesy and Cartography of the State Geological Committee of the Soviet Union.
Brunetti, A., Weder, B., 2003. A free press is bad news for corruption. J. Public Econ. 87 (7–8), 1801–1824.
Buonanno, P., Vanin, P., 2017. Social closure, surnames and crime. J. Econ. Behav. Organ. 137, 160–175.
Campos, J.E., Lien, D., Pradhan, S., 1999. The impact of corruption on investment: predictability matters. World Dev. 27 (6), 1059–1067.
Cavalli-Sforza, L.L., Moroni, A., Zei, G., 2004. Consanguinity, inbreeding, and genetic drift in italy. No. 39 in Monographs in Population Biology. Princeton University Press, Princeton.
Charron, N., Dijkstra, L., Lapuente, V., 2014. Regional governance matters: quality of government within european union member states. Reg. Stu. 48 (1), 68–90.
Charron, N., Dijkstra, L., Lapuente, V., 2015. Mapping the regional divide in europe: a measure for assessing quality of government in 206 european regions. Soc. Indic. Res. 122 (2), 315–346.
Chuah, S.H., Fahoum, R., Hoffmann, R., 2013. Fractionalization and trust in india: a field-experiment. Econ. Lett. 119 (2), 191–194.
CIA, E., 1996. The World Factbook (published online). Intelligence Agency, Washington, DC.
Conley, T.G., 1999. Gmm estimation with cross sectional dependence. J. Econ. 92 (1), 1–45.
Courbage, Y., Todd, E., 2014. A convergence of Civilizations: The Transformation of Muslim Societies Around the World. Columbia University Press.
Del Monte, A., Papagni, E., 2007. The determinants of corruption in Italy: regional panel data analysis. Eur. J. Political Econ. 23 (2), 379–396.
Di Liberto, A., Sideri, M., 2015. Past dominations, current institutions and the italian regional economic performance. Eur. J. Political Econ. 38, 12–41.
Dincer, O.C., 2008. Ethnic and religious diversity and corruption. Econ. Lett. 99 (1), 98–102.
Duranton, G., Rodríguez-Pose, A., Sandall, R., 2009. Family types and the persistence of regional disparities in europe. Econ. Geogr. 85 (1), 23–47.
Easterly, W., Levine, R., 1997. Africa's growth tragedy: policies and ethnic divisions. Q. J. Econ. 1203–1250.
Edlund, L., 2018. Cousin marriage is not choice: muslim marriage and underdevelopment. AEA Papers Proc. 108. 353–57
Ekelund, R.B., Tollison, R.D., Anderson, G.M., Robert, F., Davidson, A.B., et al., 1996. Sacred Trust: The Medieval Church as an Economic Firm. Oxford University Press.
Elbahnasawy, N.G., Revier, C.F., 2012. The determinants of corruption: cross-country-panel-data analysis. Dev. Econ. 50 (4), 311–333.
Enke, B., 2019. Kinship, cooperation, and the evolution of moral systems. Quart. J Econ. 134 (2), 953–1019.
Ermisch, J., Gambetta, D., 2010. Do strong family ties inhibit trust? J. Econ. Behav. Organ. 75 (3), 365–376.
Fiorino, N., Galli, E., Petrarca, I., 2012. Corruption and growth: evidence from the italian regions. Eur. J. Gov. Econ. 1 (2), 126–144.
Fisman, R., Gatti, R., 2002. Decentralization and corruption: evidence across countries. J. Public Econ. 83 (3), 325–345.
Francois, P., Zabojnik, J., 2005. Trust, social capital and economic development. J. Eur. Econ. Assoc. 3 (1), 51–94.
Gallup, J.L., Sachs, J.D., Mellinger, A.D., 1999. Geography and economic development. Int. Reg. Sci. Rev. 22 (2), 179–232.

Glaeser, E.L., Saks, R.E., 2006. Corruption in america. J. Public Econ. 90 (6), 1053–1072.
Glatzer, B., 2002. The Pashtun Tribal System. Concept of Tribal Society, pp. 265–282.
Golden, M.A., Picci, L., 2005. Proposal for a new measure of corruption, illustrated with italian data. Econ. Politics 17 (1), 37–75.
Goody, J., 1983. The Development of the Family and Marriage in Europe. Cambridge University Press.
Gounev, P., Bezlov, T. 2010. Examining the links between organised crime and corruption. CSD.
Greif, A., 2006. Family structure, institutions, and growth: the origins and implications of western corporations. Am. Econ. Rev. 308–312.
Greif, A., Tabellini, G., 2015. The clan and the corporation: sustaining cooperation in China and Europe. Available at SSRN 2565120.
Gunnemark, E., 1991. Countries, Peoples, and Their Languages: The Geolinguistic Handbook. Summer Institute of Linguistics, Inc., Dallas, Texas.
Hamilton, W.D., 1964. The genetical evolution of social behaviour. I. J. Theor. Biology 7 (1), 1–16.
Hamilton, W.D., 1975. Innate social aptitudes of man: an approach from evolutionary genetics. Biosocial Anthropol. 133, 155.
Hammarström, H., Nordhoff, S., 2011. Langdoc: bibliographic infrastructure for linguistic typology. Oslo Stud. Lang. 3 (2), 31–43.
Henrich, J., 2015. The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter. Princeton University Press, Princeton.
Herlihy, D., 1985. Medieval Households. Harvard University Press.
Hoben, A.D., Buunk, A.P., Fincher, C.L., Thornhill, R., Schaller, M., 2010. On the adaptive origins and maladaptive consequences of human inbreeding: parasite prevalence, immune functioning, and consanguineous marriage. Evol. Psychol. 8 (4), 658–676.
Hofstede, G.H., 2001. Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations. Sage, New York.
Howell, L.D., 2011. International Country Risk Guide Methodology. PRS Group, East Syracuse, NY.
Kaufmann, D., Kraay, A., Mastruzzi, M., et al., 2010. The worldwide governance indicators: methodology and analytical issues. Tech. Rep., The W Bank.
Kirby, K.R., Gray, R.D., Greenhill, S.J., Jordan, F.M., Gomes-Ng, S., Bibiko, H.J., Blasi, D.E., Botero, C.A., Bowern, C., Ember, C.R., et al., 2016. D-Place: a global database of cultural, linguistic and environmental diversity. PLoS One 11 (7), e0158391.
Korotayev, A., Kazankov, A., Borinskaya, S., Khaltourina, D., Bondarenko, D., 2004. Ethnographic atlas xxx: peoples of siberia. Ethnology 43 (1), 83–92.
Korotayev, A.V., 2003. Unilineal descent organization and deep christianization: a cross-cultural comparison. Cross-Cultur. Res. 37 (1), 133–157.
Kugler, M., Verdier, T., Zenou, Y., 2005. Organized crime, corruption and punishment. J. Public Econ. 89 (9), 1639–1663.
La Porta, R., Shleifer, A., Vishny, R., 1998. Law and finance. J. Political Econ. 106, 1113–1155.
La Porta, R., Lopez-de Silanes, F., Shleifer, A., Vishny, R., 1999. The quality of government. J. Law Econ. Organ. 15 (1), 222–279.
Lambsdorff, J.G., 2003. How corruption affects persistent capital flows. Econ. Gov. 4 (3), 229–243.
American Association of Law Libraries, A., Flores, A.A., 1989. Foreign law: Current sources of codes and basic legislation in jurisdictions of the world. Littleton, Colo.: FB Rothman.
Lewis, A., 2014. Security, Clans and Tribes: Unstable Governance in Somaliland, Yemen and the Gulf of Aden. Palgrave Macmillan, London.
Mauro, P., 1995. Corruption and growth. Q. J. Econ. 681–712.
Mauro, P., 1998. Corruption and the composition of government expenditure. J. Public Econ. 69 (2), 263–279.
Michalopoulos, S., 2012. The origins of ethnolinguistic diversity. Am. Econ. Rev. 102 (4), 1508.
Mitterauer, M., 2010. Why Europe?: The Medieval Origins of Its Special Path. University of Chicago Press.
Moscona, J., Nunn, N., Robinson, J.A., 2017. Keeping it in the family: Lineage organization and the scope of trust in sub-Saharan Africa American Economic Review 107, 565–571.
Moscona, J., Nunn, N., Robinson, J. A., 2018. *Social structure and conflict: Evidence from sub-Saharan Africa* (No. w24209). National Bureau of Economic Research.
Muller, S.H., 1964. The World's Living Languages: Basic Facts of Their Structure, Kinship, Location, and Number of Speakers. Ungar.
Murdock, G.P., 1962–1971. Ethnographic atlas, installments i-xxvii. In: Ethnology, pp. 1–10.
N., U., 1995. Demographic Yearbook. United Nations, New York.
New, M., Lister, D., Hulme, M., Makin, I., 2002. A high-resolution data set of surface climate over global land areas. Climate Res. 21 (1), 1–25.
Nordhaus, W.D., 2006. Geography and macroeconomics: new data and new findings. Proc. Natl. Acad. Sci. United States Am. 103 (10), 3510–3517.
Nunn, N., Puga, D., 2012. Ruggedness: the blessing of bad geography in africa. Rev. Econ. Stat. 94 (1), 20–36.
Persson, T., Tabellini, G., Trebbi, F., 2003. Electoral rules and corruption. J. Eur. Econ. Assoc. 1 (4), 958–989.
Putnam, R.D., Leonardi, R., Nanetti, R.Y., 1994. Making Democracy Work: Civic Traditions in Modern Italy. Princeton University Press, Princeton.
Ramachandran, S., Deshpande, O., Roseman, C.C., Rosenberg, N.A., Feldman, M.W., Cavalli-Sforza, L.L., 2005. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in africa. Proc. Natl. Acad. Sci. United States Am. 102 (44), 15942–15947.
Ramankutty, N., Foley, J.A., Norman, J., McSweeney, K., 2002a. The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. Global Ecol. Biogeogr. 11 (5), 377–392.
Ramankutty, N., Foley, J.A., Norman, J., McSweeney, K., 2002b. The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. Global Ecol. Biogeogr. 11 (5), 377–392.
Roberts, J., 1962. Sociocultural change and communication problems. Study of the Role of Second Languages in Asia, Africa, and Latin America. pp. 105–23
Schulz, J., Bahrami-Rad, D., Beauchamp, J., Henrich, J. The Origins of WEIRD Psychology (June 22, 2018). Available at SSRN: https://ssrn.com/abstract=3201031 or http://dx.doi.org/10.2139/ssrn.3201031.
Schulz, J. The Churches' Bans on Consanguineous Marriages, Kin-Networks and Democracy (June 12, 2017). Available at SSRN: https://ssrn.com/abstract=2877828 or http://dx.doi.org/10.2139/ssrn.2877828.
Serra, D., 2006. Empirical determinants of corruption: a sensitivity analysis. Public Choice 126 (1–2), 225–256.
Tabellini, G., 2010. Culture and Institutions: Economic Development in the Regions of Europe. J. Eur. Econ. Assoc. 8 (4), 677–716.
Tanzi, V., Davoodi, H.R., 1998. Roads to nowhere: how corruption in public investment hurts growth, vol. 12. International Monetary Fund.
Todd, E., 1990. L'invention de l'Europe [The invention of Europe]. Seuil.
Todd, E., Garrioch, D., 1985. The explanation of ideology: Family structures and social systems. Blackwell Oxford.
Treisman, D., 2000. The causes of corruption: a cross-national study. J. Public Econ. 76 (3), 399–457.
Uslaner, E.M., 2005. Trust and corruption. In: Lambsdorff, J.G., Taube, M., Schramm, M. (Eds.), The new institutional economics of corruption. Routledge, pp. 76–93.
Walker, R.S., Bailey, D.H., 2014. Marrying kin in small-scale societies. Am. J. Human Biol. 26 (3), 384–388.
Weber, M., 1958. The protestant ethic and the spirit of capitalism, trans. In: Talcott parsons (new york: Charles Scribner's Sons, 1958, p. 182.
Woodley, M.A., Bell, E., 2013. Consanguinity as a major predictor of levels of democracy: a study of 70 nations. J. Cross-Cultural Psychol. 44 (2), 263–280.
Worldmark, 1995. Worldmark encyclopedia of the nations. Gale Cengage.
Yamagishi, T., Cook, K.S., Watabe, M., 1998. Uncertainty, trust, and commitment formation in the united states and japan 1. Am. J. Sociol. 104 (1), 165–194. AJSv104
Yamagishi, T., Yamagishi, M., 1994. Trust and commitment in the united states and japan. Motivation Emotion 18 (2), 129–166.