

# **The Role of Geography and History in Cultural Diffusion: A Gravitational Approach**

Pauline Grosjean

*University of California at Berkeley*

Nov 10, 2008

*Preliminary and incomplete version*

## **Abstract**

By using micro-level data from a household survey in 28 European and central Asian transition countries and Turkey, this paper develops a gravitational approach that allows to empirically investigate the respective roles of local social interactions, proxied by geographic localization data, and of distant history, in particular the influence of former dynastic empires, on the diffusion of opinions and on some economic outcomes. The focus is on opinions and economic outcomes for which both vertical and horizontal transmission have been shown or conjectured to play a role, such as: preferences for the welfare state, general trust, corruption, economic occupation and female labor force participation. The theoretical model predicts that cultural distance and dissimilarity of economic outcome will increase in physical distance. This is confirmed by empirical results. Other results show evidence of significant negative “border effects” on cultural and economic dissimilarity, and a strong negative effect of past empires, as well as of EU integration and internet penetration. The results also indicate that Turkey is quite dissimilar to the new Europe, but not fundamentally different from other EU candidates.

Keywords: culture, economic development, Europe, geography, gravity model, Turkey.

JEL classification: F59 ; N50 ; O10 ; P30.

## 1. Introduction

What are the respective roles of geography and history in the formation of attitudes and in the determination of economic outcomes? The empirical influence of, on the one hand, geography, through local social interactions and, on the other hand, distant history, through the vertical transmission of beliefs (Tabellini, 2007) or the long term persistence of institutions (Acemoglu et al. 2001), on cultural values and economic outcomes have generally been investigated separately in the literature. However, the two may be largely intertwined. For example, geographical factors may contribute to explain the observed divergence between the former Habsburg empire, whose geographical localization facilitated exposure to Enlightenment values, and the former Ottoman Empire, who, under the constant threats of Mongol and Russian invasions, had to organize itself as a military state, which had negative subsequent consequences for its economic development (Dimitrova-Grajzl 2008). In a recent paper, Guiso, Sapienza and Zingales (2008) observe that neither of the two main competing transmission mechanisms explaining the long term persistence of economic development: the institution transmission mechanism of Acemoglu et al. (2001) and the cultural transmission mechanism of Tabellini (2007) can fully reject the “alternative” (or rather – complementary) view that the source of persistence is geographical (Guiso, Sapienza and Zingales, 2008). Despite the fact that theoretical models of cultural transmission (Bisin Verdier, 2001) take explicitly into account the influence of social interactions (“oblique” transmission mechanisms) on “vertical” parental transmission in the formation of cultural values or economic outcomes, such as corruption (Hauk and Saez Marti 2001), most empirical studies of vertical transmission, such as Tabellini (2008) or Fernandez (2006) abstract of such social interactions<sup>1</sup>. This paper attempts at filling this gap in the literature, by modeling jointly the influence of historical and geographic factors.

The role of local social interactions on culture and economic outcome has been well established in the literature. Critical mass models, where the utility of an action is positively related to the number of agents who already have taken the action, or local social interaction models have been applied, for example, to the persistence of income inequality (Durlauf, 1996), crime (Glaeser et al., 1995; Sah, 1991), saving norms (Lindbeck, 1997), the transmission of culture (Bisin and Verdier, 2001) or economic occupation (Topa, 2001). Bertrand et al. (2000) develop an empirical

---

<sup>1</sup> Despite fact that this geographica componenet is modeled in the theortica part of Atbellinin 2008.

application and show the presence of strong network effects in welfare state participation. They argue that this effect is due to the propagation of information among social network. An alternative explanation could be that it is related to the social acceptance of welfare use among the network. For instance, Fernandez (2006) argue that social acceptability of female work among peer groups has played an important role in the rise of female labor force participation.

The recent surge of interest in the role of culture in economics has also well established the influence of history on culture and economic development. Many empirical studies have shown how processes of “vertical” or intergenerational transmission of beliefs and preferences can explain the persistence of inefficient institutions (Tabellini 2007, 2008, Dimitrova and Grajzl 2008), the formation of social capital (Guiso, Sapienza and Zingales, 2007) unemployment (Brugger et al., 2008) and contribute to the understanding of the microeconomic behavior of agents in their fertility decision (Fernandez), occupational choice (Botticini and Epstein 2006a 2006b), willingness to trust and to cooperate with others (Tabellini, 2007, 2008) or attitudes towards the welfare state (Alesina and Fuchs Schundeln, 2008). The main idea of the underlying theoretical literature is that the backward looking behavior of parents who try and transmit their own values to their children generate some hysteresis that can explain the slow moving character of institutions (Roland, 2004) and the slow adaptation of work or fertility decisions to new economic environment (Fernandez). In a seminal paper, Bisin and Verdier (2001) endogenize preference traits (or “culture”) in a process of intergenerational transmission of traits. They model how preferences of children are acquired through an adaptation and imitation process which depends on their parents’ socialization actions, and on the characteristics of the economic and social environment in which they later evolve. Hauk and Saez-Marti (2001) develop a similar model and show how level of corruption is determined by the effort of parents in transmitting values of honesty, which itself depends on the distribution of ethics in society. Tabellini (2008) describes a model where individuals cooperative behavior is determined both by economic incentives and by norms of good conduct inherited from earlier generations. In all of these models, parental transmission mechanisms are explicitly influenced by the diffusion of norms and behaviors in society: to keep the epidemiological analogy, “vertical” and “horizontal” transmission mechanisms (see Cavalli Sforza) are substitute. This theoretical literature hence bridges the gap between the social interaction and the culture and economics empirical literature. However, a limited number of empirical studies have managed to reconcile the two approaches.

By using micro-level data, we propose a gravitational approach that allows to empirically investigate the respective roles of geography and of history on the diffusion of opinions and on some economic outcomes. We focus on opinions and economic outcomes for which both vertical and horizontal transmission have been shown or conjectured to play a role, such as: preferences for the welfare state, general trust, corruption, economic occupation and female labor force participation. Indeed, Alesina and Fuchs Schundeln (2007) study the influence of past history on fairness and attitudes towards the welfare state, while Lindbeck and al (2003) study the influence of work norms on voters' preferences for the welfare state and Bertrand et al. (2000) study the influence of local social interactions on welfare use. Tabellini (2007, 2008) discuss how distant past history, but also local social interactions (see the theoretical model in Tabellini 2008) act together to determine how much people trust each other (general trust). Hauk and Saez Marti (2001) show how level of corruption is determined by inherited values of honesty which parents try to transmit to their offspring, but also by the prevalence of ethics in society and the economic payoffs of corruption. Scheinkman and de Paula (2008) study spatial patterns of corruption, and show that the informality of firms is correlated with the informality of firms from which it buys or sell. Botticini and Eckstein (2006b) discuss how economic occupations are path dependent and influenced by changes in distant history, while many social network models since the pioneering work of Granovetter (1973) and economic geography models put forward the role of location in the determination of occupation patterns. Similar arguments can be made in the case of female labor force participation, and more specifically, Fernandez (2006, etc) focus on the role of inherited values, the prevalence of female work in the social environment, and the perception of female work by peer groups and neighbors as determinant of female labor force participation.

The theoretical model in Section 2 of this paper extends previous theoretical models of cultural transmission by specifying that the social interaction which influences vertical transmission mechanism is geographical in nature. Such models usually assume that preferences of children are acquired through an adaptation process which depends on the parents' socialization actions, *and* on the social environment in which children live. However, the nature of social interaction is not precisely specified: the socialization effort of parents itself depends on the distribution of cultural traits in society as a whole. This assumption is hardly realistic, not the least because of information problems. Instead, we consider that the relevant social environment to which parents

refer to is their peer group, which, following Glaeser, Sacerdote and Sheinkman (1996), is assumed to consist of geographical neighbors. We test the predictions of this theoretical model using a gravity model, partly in order to avoid Manski's (1993) reflection bias.

Our gravitational model proxies geography and horizontal cultural diffusion processes by physical distances and geopolitics characteristics, such as the presence of borders. The role of history and vertical cultural diffusion processes is captured by the distant past of each locality in our sample, namely the inclusion in former empires such as the Ottoman, Habsburg, Russian, or Prussian Empires and the length of time the locality was subjected to such influence. We are able to trace back, at the sub-national level, the belonging of each locality (Primary Sampling Unit, or PSU) to former Empires and the length of time PSUs belonged to different empires. We are thus able to identify the respective roles of a common past history, though former integration to Empires, a common history today, through integration in nation states or the European Union, and geographic characteristics, namely physical distance. Nevertheless, our results are subject to the caveat that history itself was partly determined by geography.

Gravitational models have almost exclusively been applied to trade models, with two notable exceptions. Kelejian, Murrell and Shepotyloe (2007) develop a spatial analysis of spillover between countries in the development of institutions, but use a spatial lag model instead of a gravity model. Sachs, Rodrik examine role of geography on institutions. Head Mayer (2008) use a gravity model to detect local interactions from the spatial pattern of first names in France.

The main contributions of the paper is thus to model and estimate jointly the roles of vertical and horizontal transmission processes in the determination of values and economic outcomes, by means of one of the few applications of gravity models outside of the trade literature. Thereby, this paper also contributes to the literature on border effects. The trade literature has found evidence for negative and economically very significant effect of geopolitical borders on trade. We find, on the contrary, negative and significant effects of borders on cultural and economic dissimilarity, indicating that non market interactions effects and cultural differences, which have been blamed for large negative border effects on trade, may have been overestimated. Our results, which take into account precise locations rather than country average as usual on the trade literature, show instead a continuous process of diffusion over geopolitical borders.

Section 2 presents a theoretical model of cultural diffusion in which oblique transmission mechanisms depend on the geographical localization of agents, and the econometric specification. Section 3 discusses the data. Section 4 presents the results. In Section 5, the cultural gravity model is extended to account for the impact of EU integration, with a special focus on the case of Turkey, and to study the effect of internet on cultural homogenization. Section 6 concludes.

## 2. Theoretical Model

In this section, we study a model of cultural evolution that builds on Bisin Verdier (2001) and Hauk Saez-Marti (2001). Such models consider both the influence of cultural parental transmission and social interaction. Preferences of children are indeed acquired through an adaptation process which depends on the parents' socialization actions, *and* on the social environment in which children live. However, the nature of social interaction is not precisely specified: the socialization effort of parents itself depends on the distribution of cultural traits in society as a whole. This assumption is hardly realistic, not the least because of information problems. Instead, we consider that the relevant social environment to which parents refer to is their peer group, which is, as in Glaeser, Sacerdote and Sheinkman (1996), consists of their geographical neighbors.

We consider a two-cultural trait population of individuals. The transmission of cultural traits is modeled as a mechanism which interacts “vertical” socialization inside the family and “horizontal” socialization outside the family, via imitation of role models that the children interacts with. Such social interaction is determined geographically (see also Tabellini 2007).

Suppose there are two possible cultural traits  $\{a, b\}$ . Such cultural trait can reflect preferences with regards to female labor force participation, honesty and the inclination to pay bribes, or fairness and preferences for redistribution. Individuals live in different locations, or neighborhoods, indexed by  $l$ . There are  $L$  possible locations. Localities neighboring to locality  $l$  are denoted by  $C(l)$ . Individuals in each locality are heterogeneous in two different ways.

Firstly, as described above, they may differ in their cultural trait. Secondly, they are member of larger groups with different mean preferences (as in Head and Mayer, 2007). In each locality, there are  $G$  groups, indexed by subscript  $g$ . The share of each group in each location is given by:  $x_{lg}$  with  $\sum_{g=1}^G x_{lg} = 1$ . The portion of each group with cultural trait  $a$  is denoted by  $s_{lga}$ . The prevalence of cultural trait  $a$  in neighborhood  $l$  is then:  $s_{la} = \sum_{g=1}^G s_{lga}x_{lg}$ . The prevalence of cultural trait  $b$  is:  $s_{lb} = 1 - s_{la}$ .

Following Hauk and Saez-Marti (2001), we consider an overlapping generation model, with a Poisson birth and death process. With probability  $\gamma$ , an active agent is active the next period. With probability  $(1 - \gamma)$ , an active agent in  $t$  has a child, which becomes active in  $t+1$ . Children are born without predetermined preferences or cultural trait. As in Bisin and Verdier (2001) and Hauk and Saez-Marti (2001), parents' education effort  $\tau$  corresponds to the probability of successful "vertical" transmission of cultural traits, that is to say the probability that the child ends up with identical preferences as his parent. If vertical socialization is unsuccessful, the child remains naïve and gets randomly matched with somebody else, whose cultural trait he adopts. The innovation of our model is that this match depends on the geographical localization of agents. As in Tabellini (2007), we consider that the probability of a match decreases with the distance that separates two agents. More precisely, we consider that agents are either matched with an individual within their neighborhood or with an individual from an adjacent location  $n$  with a probability  $\vartheta_{tln}$  that reflects the nature of social interactions, and that decreases with distance.  $\vartheta_{tln}$  is assumed to vary between zero<sup>2</sup> (localized interactions) and  $1/(L - 1)$  (globalized interactions) (Head and Mayer, 2007). Each household in each of the  $G$  groups socializes his children according to this socialization process, but each child only interacts with people of her parents' group.

Let  $p^{ij}$  denote the probability that a child from a family with trait  $i$  is of type  $j$  and  $\tau^i$  the education effort by the parent of type  $i$ . The socialization mechanism described above implies that:

$$p_t^{aa} = \tau_t^a + (1 - \tau_t^a)(s_{lga} + \sum_{n \neq l} \vartheta_{tln} s_{ngat}) \quad (1)$$

---

<sup>2</sup> This brings us back to the models in Bisin and Verdier (2001) and Hauk and Saez-Marti (2001).

$$p_t^{ab} = (1 - \tau_t^a)[1 - (s_{l gat} + \sum_{n \neq l} \vartheta_{tln} s_{ngat})] \quad (2)$$

Where  $s_{k gat}$ ,  $k \in \{l, n\}$  is the proportion of active agents of type  $a$  in locality  $k$ . Similarly, we get:

$$p_t^{bb} = \tau_t^b + (1 - \tau_t^b)[1 - (s_{l gat} + \sum_{n \neq l} \vartheta_{tln} s_{ngat})] \quad (3)$$

$$p_t^{ba} = (1 - \tau_t^b)(s_{l gat} + \sum_{n \neq l} \vartheta_{tln} s_{ngat}) \quad (4)$$

As in Bisin and Verdier (2001), parents display imperfect empathy. When taking the decision to socialize their children, they maximize their children's utility, but they evaluate such welfare through their own preferences. In addition, we assume that utility of a child depends on what group  $g$  he belongs to<sup>3</sup>. Let  $V_g^{ij}$  denotes the utility to a type  $i$  parent belonging to group  $g$  of a type  $j$  child,  $i, j \in \{a, b\}$ . In order to determine the optimal socialization effort  $\tau_g$ , which costs  $C(\tau_g)$ , parents maximize:

$$p_t^{ii} V_g^{ii} + p_t^{ij} V_g^{ij} - C(\tau_{gt})$$

We obtain the following optimal education efforts:

$$C'(\tau_g^a) = (V_g^{aa} - V_g^{ab})[1 - (s_{l gat} + \sum_{n \neq l} \vartheta_{tln} s_{ngat})] \quad (5)$$

$$C'(\tau_g^b) = (V_g^{bb} - V_g^{ba})(s_{l gat} + \sum_{n \neq l} \vartheta_{tln} s_{ngat}) \quad (6)$$

We can now characterize the dynamic behavior of  $s_{la}$ :

$$s_{l gat+1} = \gamma s_{l gat} + (1 - \gamma)[s_{l gat} p_t^{aa} + (1 - s_{l gat}) p_t^{ba}]$$

Substituting (1)-(4) and rearranging, we obtain:

$$s_{l gat+1} = s_{l gat} + (1 - \gamma)\{s_{l gat}(\tau_g^a - \tau_g^b)(1 - s_{l gat} - \sum_{n \neq l} \vartheta_{tln} s_{ngat}) + (1 - \tau_g^b) \sum_{n \neq l} \vartheta_{tln} s_{ngat}\} \quad (7)$$

---

<sup>3</sup> We can think of these groups as socio economic groups, or religious groups.

Under stationary expectations ( $V_{gt}^{aa} - V_{gt}^{ab} = V_g^{aa} - V_g^{ab}$  and  $V_t^{bb} - V_t^{ba} = V^{bb} - V^{ba}$  for all  $t$ ),

Hauk and Saez-Marti (2001) show that (5) has three rest points, (i)  $s_{lga} = 0$ , (ii)  $s_{lga} = 1$ , and an interior rest point. In this modified model, the interior rest point is given by:

$$s_{lga}^* = \frac{V_g^{aa} - V_g^{ab}}{V_g^{bb} - V_g^{ba} + V_g^{aa} - V_g^{ab}} - \sum_{n \neq l} \vartheta_{ln} s_{nga}^* \quad (8)$$

with  $\tau_g^a(s_{lga}, V_g^{aa} - V_g^{ab}) = \tau_g^b(s_{lga}, V_g^{bb} - V_g^{ba})$

We can see from (5) and (6) that vertical cultural transmission (parents' effort) and horizontal transmission (from peer groups) are substitutes. Under this condition, the interior rest point is globally stable, while the other rest points are unstable (Bisin and Verdier 2001).

Let us denote  $c_g = \frac{V_g^{aa} - V_g^{ab}}{V_g^{bb} - V_g^{ba} + V_g^{aa} - V_g^{ab}}$  and drop the subscript  $a$  in what follows.

The total share of type  $a$  in locality  $l$  is given by  $s_l = \sum_{g=1}^G s_{lg} x_{lg}$ , which, using (8) can be rewritten as:

$$s_l^* = \sum_{g=1}^G c_g x_{lg} - \sum_{g=1}^G x_{lg} \sum_{n \neq l} \vartheta_{ln} s_{ng}^* \quad (9)$$

Replacing by the value of  $s_{ng}^*$ , which can be obtained in a similar way as (8), we obtain:

$$s_l^* = \sum_{g=1}^G c_g x_{lg} - \sum_{g=1}^G c_g x_{lg} \sum_{n \neq l} \vartheta_{ln} + A_l \quad (10)$$

With  $A_l = \sum_g x_{lg} \sum_{n \neq l} \vartheta_{ln} \sum_{m \neq n} \vartheta_{nm} s_{mga}$  and  $m \in C(n)$ .

This expression shows how the share of cultural trait in one locality depends on the share of the trait in the adjacent localities, as well as in all the localities adjacent to it, and so on. The share of cultural trait in one neighborhood depends on this way on the share of the trait in all neighborhoods that are connected with each other.

Estimating equation of type (10) would run into severe complications and shortfall. We would need to specify each neighbor of each localities and so on. Although possible and done in the literature before (see Case, ), this would be a daunting task given the size of the sample we have at hands. Furthermore, as described by Manski (1993) the OLS estimates of this equation would

be biased. The prevalence of the cultural trait  $a$  in the adjacent location  $n$  depends in part on the share of the cultural trait in location  $l$ . This is what Manski labels the “reflection” problem.

Instead, we are more interested in the dissimilarity in prevalence of cultural traits between any pair of locations:  $|s_l - s_n|$ . As explained in Head and Mayer (2007), incorporating  $s_n$  in the left hand side variable addresses the reflection problem. Let us drop the subscript  $a$  to make things more tractable.

Following steps (8) and (9), we can define the total share of type  $a$  in locality  $n$ :

$$s_n^* = \sum_{g=1}^G c_g x_{ng} - \sum_{g=1}^G x_{ng} \sum_{m \neq n} \vartheta_{mn} s_{mg}^*$$

Adopting vector notations:  $\sum_{g=1}^G c_g x_{lg} = \beta \cdot x_l$ , and denoting  $A_n = \sum_{g=1}^G x_{ng} \sum_{m \neq n} \vartheta_{mn} s_{mg}^*$  we obtain the following expression for the dissimilarity in the prevalence of cultural trait  $a$  among a pair of locations:

$$|s_l - s_n| = |\beta \cdot (x_l - x_n) - \beta \cdot x_l \sum_{n \neq l} \vartheta_{ln} + A_l - A_n| \quad (8)$$

This equation has three main implications that we can investigate empirically.

Firstly, similar group composition entails more similarity in the prevalence of cultural trait between two locations.

Secondly, distance increases cultural dissimilarity between locations. Indeed,  $|s_l - s_n|$  decreases in  $\vartheta_{ln}$ , which reflects the extent of spatial interaction and decreases with distance.

Prevalence of cultural traits in third locations’ ( $A_l - A_n$ ) influence bilateral dissimilarity. The empirical estimation will attempt to control these by using fixed effects.

We propose the following linear estimation approach that accounts for the properties of the model:

$$MDC_{ln} = \alpha_1 MDS_{ln} + \alpha_2 MDE_{ln} + \alpha_3 MDO_{ln} + \alpha_4 MDR_{ln} + \alpha_5 DIST_{ln} + \alpha_6 I + \alpha_7 C + \alpha_8 E + \delta_i + \delta_j + \epsilon_{ln} \quad (9)$$

Where the MD variables are metrics of dissimilarity between pairs of locations  $l$  and  $n$ . In each case, following Head and Mayer (2007), we use the Manhattan Distance which sums over the absolute differences in shares of responses for multinomial variables. For example, MDC in the case of general trust is defined as:  $MDC_{ln} = \sum_{i=1}^I |s_{il} - s_{in}|$  where  $s_{il}$  (respectively  $s_{in}$ ) is the share, in location  $l$  (respectively  $n$ ) of responses allocated to each modality  $i$  of the  $I$  modalities of the following WVS question on general trust that has been used also in this survey: “Generally Speaking, would you say that most people can be trusted, or that you can’t be too careful in dealing with people?” Responses modalities are on a scale of 1 to 5, with 1 corresponding to “complete distrust and 5 to “Complete Trust”. Dissimilarities across localities for each of the variables of interest (See next Section) are defined in a similar way. For dummy variable: the Manhattan Distance is just the absolute value of the difference between two locations.

The covariates in (9): Social class dissimilarity  $MDS$ , education dissimilarity  $MDE$ , occupation dissimilarity  $MDO$  and religious affiliation dissimilarity  $MDR$  are defined in a similar way.

The next covariate in (9) examines the role of geographic distance between two locations. We expect  $\alpha_5$  to be positive.  $I$  is a dummy variable indicating whether the two districts are located in the same country. As usual in the gravity trade models,  $C$  is a dummy variable to indicate whether the two countries in which each location belongs are contiguous. The role of history is investigated through the use of a dummy  $E$  indicating whether two locations belonged to the same Empire in the past. The following section describes in more details the definitions of the different Empires whose role we examine. We also use a different specifications where we investigate the impact of the length of time during which two locations belonged to the same Empire. Finally, the regression specification includes an error term and a set of intercepts  $\delta_i, \delta_j$  for each country.

### **3. Data and Preliminary Evidence on the Role of Geography and History**

#### ***3.1.Data and Dependent Variables***

- Data

The data comes from the Life in Transition Survey (LITS), a survey conducted by the European Bank for Reconstruction and Development and the World Bank in 2006, in 28 post-transition countries and Turkey<sup>4</sup>. Respondents to the survey were drawn randomly, using a two stage sampling method, with census enumeration areas as Primary Sampling Units (PSUs), and households as secondary sampling units. The definition of locations in this paper corresponds to the PSUs. The data set contains 50 PSUs in each of the 29 countries. There are thus 1450 PSU pairs for which we have observations.

The geographical distance between each pair of PSU is computed from longitudinal and latitudinal coordinates of the geographical center of the PSU using the great circle formula. For each PSU pair, we calculate our measures of cultural distance and of covariate dissimilarity defined in (9). The analysis is then conducted on the cross section of 1450<sup>2</sup> distance measures.

We use the Periodical Atlas of Europe in order to reconstruct Empire delimitations and their evolution across time, from 1300 to 2000. 1300 corresponds to the start of empire consolidation in Medieval Europe and is thus taken as the starting date of this study (Dimitrova-Grajzl 2008). Table A1 presents some descriptive statistics on the length and geographical delimitation of all the empires whose impact we investigate.

- Dependent variables

Preferences for redistribution and the welfare state:

Preferences for redistribution and the welfare state are assessed using the following survey question:

*Do you think the state should be involved in the following? (Scale: not involved at all/let the market play =1; Somewhat/moderately involved =2; strongly involved=3)*

*Reducing the gap between the rich and the poor*

Corruption:

---

<sup>4</sup> Turkmenistan was not included in the survey, neither was Kosovo. We exclude Turkey from our sample because its borders with the other countries were closed during the whole socialist period, which is in contradiction with our identification assumption.

Our variable measuring the extent of corruption of public services was built using the following survey question:

*In your opinion, how often is it necessary for people like you to have to make unofficial payments/gifts in these situations?*

*(Scale: Never=1, Seldom=2, Sometimes=3, Usually=4, Always=5)*

*Interact with the road police*

*Request official documents (e.g. passport, visa; birth or marriage certificate, land register, etc...) from authorities*

*Interact with the police on matters other than traffic and other than requesting documents*

*Go to courts for a civil matter?*

*Receive medical treatment in the public health system*

*Receive public education (university, college, vocational training)*

*Request unemployment benefits*

*Request other social security benefits*

General trust:

Here, we use in the LITS survey the widely used World Value Survey question:

*Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people? Please answer on a scale from 1 to 5, where 1 means that you have complete distrust in people, and 5 means that most people can be trusted. What would it be today? And before 1989?*

*Scale: Complete distrust=1, Some distrust=2, Neither trust nor distrust=3, Some trust=4, Complete trust=5*

### Occupation and Female labor force participation:

The categories used in the construction of the Occupation Manhattan distance are the following: unemployed; pensioner; student; housewife; or employed, with this last category broken up in different subcategories of: white collar, blue collar service worker and farmer or farm-worker. The different categories of employment were constructed from the responses about the respondent's first job, using the ISCO classification<sup>5</sup>.

Our first measure of female labor force participation corresponds to the proportion of women of working age in the PSU who are economically active (that is to say who were active in the 7 days preceding the survey or had a job to return to). Our second measure of female labor force participation aims at capturing the gender bias in the transition (or development) process since 1989. It measures the proportion of women of working age today and in 1989, who were active in 1989 and who are still active (in the above definition of activity). An alternative measure is the differential of this ratio between men and women in the PSU.

Table A2 in the appendix presents descriptive statistics of the dependent and independent variables.

### ***3.2.Preliminary Evidence on the Role of history:***

Table 1 illustrates the large differences that can be observed within empires in terms of preferences for redistribution , the extent of corruption if public services, general trust or female labor force participation.

### **Table 1: Means of Selected Preferences and Economic Outcomes in Former Empires**

---

<sup>5</sup> The ISCO categories corresponding to our white collar category are: 1: Legislator, Senior Official and Manager, 2: Professionals, and 3: Technicians and Associated Professionals. Our service workers category consists of: 4: Clerks and 5: Service workers and shop and market sales workers. 6: Skilled agricultural and fishery workers are in our "farmer and farm worker category" together with independent farmers. All the remaining ISCO categories are considered as blue collar workers.

	Preferences for redistribution	General trust	Corruption	FLP	Evolution of FLP
Ottoman Empire	0.79	3.62	1.93	0.49	0.26
Habsburg Empire	0.60	3.40	1.11	0.61	0.35
Russian Empire	0.68	3.83	1.78	0.53	0.26
Prussia	0.54	3.45	0.95	0.55	0.26
Ottoman and Habsburg	0.77	3.49	1.40	0.48	0.22
Ottoman and Russian	0.56	3.59	2.27	0.57	0.34
Habsburg and Russian	0.68	3.26	3.60	0.62	0.23
Habsburg and Prussia	0.45	3.39	0.76	0.44	0.21
Russian and Prussia	0.64	3.57	0.57	0.68	0.40

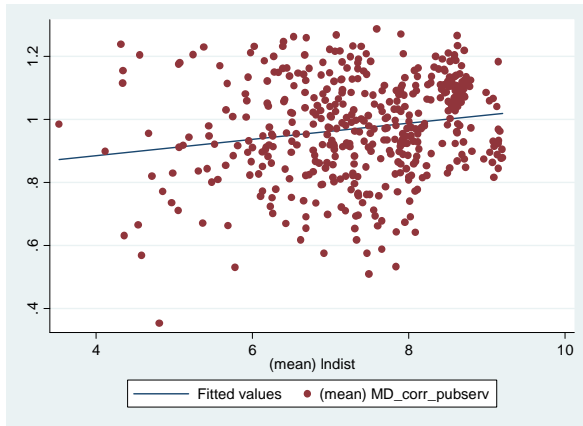
This table illustrates how past institutions may have long lasting effects. For example, preferences for redistribution and corruption are higher in the former Ottoman Empire. Historical studies reveal how the structure of the Ottoman Empire was prone to corruption. Dimitrova Grajzl cites historical studies that reveal how “Corruption became prevalent with the decline of the Ottoman Empire. “From the seventeenth century onward the typical Ottoman official holding a position of any importance regarded it as a private investment from which he was justified in deriving as large a return as possible” (Stavrianos 2000, p. 120).”. Similarly, the Ottoman Empire was quite egalitarian (Berend, 2003), which may have had long lasting effects on preferences for redistribution today. On the contrary, corruption is lowest in what was Prussia and in the Habsburg Empire, which also corresponds to historical accounts (REF) An interesting feature is that locations that belonged to different empires tend to display a mix of the values of such empires. For example, the locations that have belonged both to the Ottoman Empire and the Habsburg Empire suffer from a level of corruption that is intermediate between the levels of the former Ottoman and Habsburg Empires.

### ***3.3.Preliminary Evidence on the role of geography***

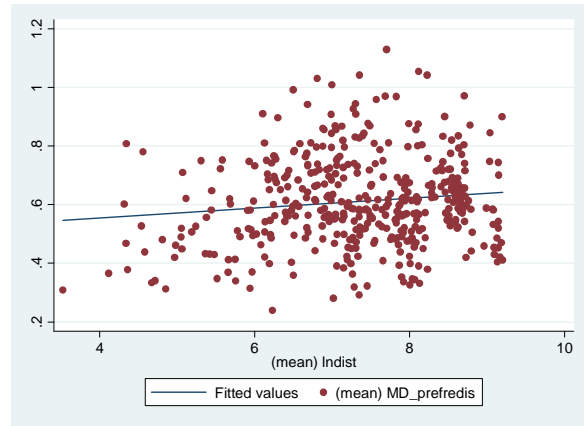
A few graphs can illustrate that cultural dissimilarity increases in distance. The following graphs plot our measure of cultural dissimilarity, the Manhattan Distance, for each of our dependent variables: preferences for redistribution, corruption, general trust, occupation and female labor force participation.

Figure 1: Manhattan Distance of Dependent Variables and Physical Distance

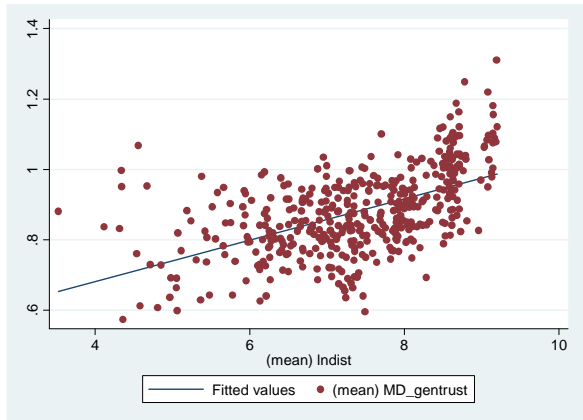
(a) Corruption:



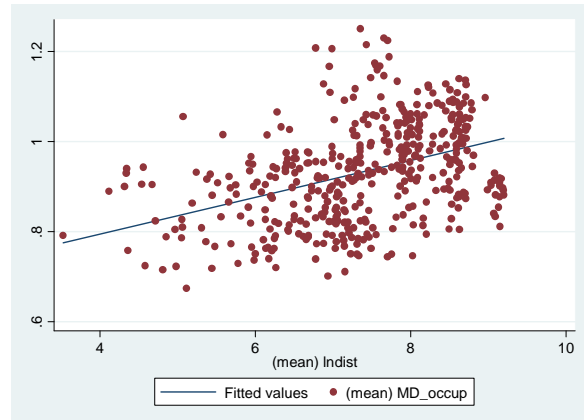
(b) Preferences for redistribution



(c) General trust:

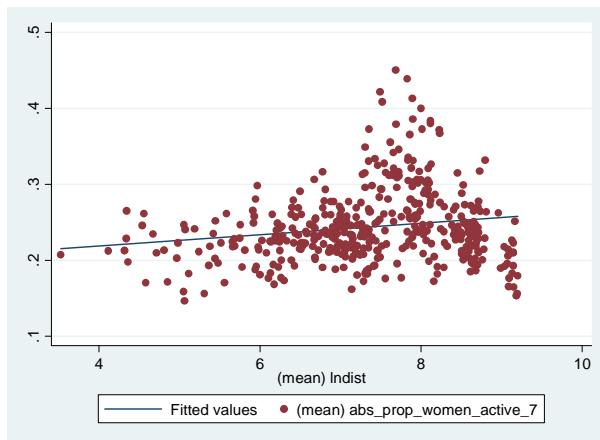


(d) Occupation:

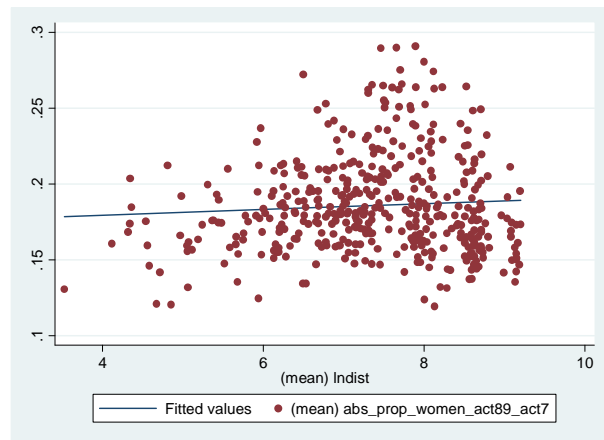


Female labor force participation:

(e) Proportion of active women



(f) Proportion of women that were active in 1989 and are still active



The influence of distance is not homogenous however. Distance seems to impact dissimilarity in terms of general trust, corruption, or economic occupation more than preferences for redistribution or female labor force participation. This may be due to the fact that the nature of social interactions may vary across different values or economic outcomes. Female labor force participation, for example, may be more influenced by “culture” or vertical transmission mechanisms, rather than by local interactions. General trust and corruption, on the contrary, are likely to be strongly influenced both by local spatial interactions and by culture. The structure of economic occupations may depend greatly on the nature of social networks (Granovetter), but less on culture and historical background. The question of the relative impacts of “vertical” and “horizontal” transmission mechanisms on different sets of values and economic outcomes is precisely at the core of this paper. We now turn to our integrated gravity equation, which allows to illustrate simultaneously the role of history and geography described in Tabellini (2007).

## 4. Results of the Gravity Model

### 4.1. Baseline results

Table 2a to 2d present estimation results of the baseline gravity equation described in (9).

**Table 2a: Results of baseline gravity equation: preferences for redistribution and general trust**

	1	2	3	4	5	6	7	8
	Pref. redistribution	Pref. redistribution	Pref. redistribution	Pref. redistribution	General trust	General trust	General trust	General trust
Distance	0.020*** [0.001]	-0.001 [0.001]	0.015*** [0.001]	-0.002*** [0.001]	0.023*** [0.001]	0.021*** [0.001]	0.022*** [0.001]	0.020*** [0.001]
Same Country		-0.108*** [0.003]		-0.102*** [0.003]		-0.011*** [0.003]		-0.009*** [0.003]
Contiguous		-0.060*** [0.002]		-0.056*** [0.002]		-0.008*** [0.001]		-0.006*** [0.001]
Same Empire			-0.037*** [0.002]	-0.025*** [0.002]			-0.011*** [0.001]	-0.010*** [0.001]
Observations	963966	963966	963966	963966	963966	963966	963966	963966
R-squared	0.084	0.086	0.085	0.086	0.07	0.07	0.07	0.07

**Table 2b: Results of baseline gravity equation: corruption and occupation**

	1	2	3	4	5	6	7	8
	Corruption	Corruption	Corruption	Corruption	Occupation	Occupation	Occupation	Occupation
Distance	0.009*** [0.001]	0.001* [0.001]	0.007*** [0.001]	0.000 [0.001]	0.016*** [0.000]	0.010*** [0.001]	0.015*** [0.000]	0.009*** [0.001]
Same Country		-0.063*** [0.003]		-0.059*** [0.003]		-0.028*** [0.002]		-0.026*** [0.002]
Contiguous		-0.003* [0.002]		-0.001 [0.002]		-0.027*** [0.001]		-0.025*** [0.001]
Same Empire			-0.019*** [0.002]	-0.016*** [0.002]			-0.015*** [0.001]	-0.010*** [0.001]
Observations	963966	963966	963966	963966	963966	963966	963966	963966
R-squared	0.1	0.101	0.101	0.101	0.231	0.231	0.231	0.231

**Table 2c: Results of baseline gravity equation: female labor force participation**

	1	2	3	4	5	6	7	8
	Proportion active women	Proportion active women	Proportion active women	Proportion active women	Proportion women active in 1989 and still	Proportion women active in 1989 and still	Proportion women active in 1989 and still	Proportion women active in 1989 and still
Distance	0.012***	0.009***	0.012***	0.009***	0.008***	0.007***	0.008***	0.007***

	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Same Country		-0.013***		-0.012***		-0.009***		-0.009***
		[0.001]		[0.001]		[0.001]		[0.001]
Contiguous		-0.010***		-0.009***		-0.003***		-0.003***
		[0.001]		[0.001]		[0.001]		[0.001]
Same Empire			-0.006***	-0.004***			-0.001***	-0.001
			[0.001]	[0.001]			[0.001]	[0.001]
Observations	957036	957036	957036	957036	957036	957036	957036	957036
R-squared	0.044	0.045	0.044	0.045	0.034	0.034	0.034	0.034

Notes to Table 2a, 2b, 2c:

All dependent variables are Manhattan Distance measures of dissimilarity between PSU pairs. Distance is expressed as the logarithm of the physical distance between PSUs in km. Regressions in Table 2a and columns 1 to 4 of Table 2b control for dissimilarity between pairs of PSUs in terms of religious affiliation, occupation, social class composition, education, age, as well as country fixed effects. Regressions in columns 5 to 8 of Table 2b and in Table 2c control for dissimilarity between pairs of PSUs in terms of religious affiliation, education, age, as well as country fixed effects.

Same Empire takes value 1 if both PSUs have belonged to the same Empire (as defined in section 4) for more than a 100 years.

\*\*\* indicates significance at 1%, \* significance at 10%. Standard errors are in brackets.

Most excluded covariates in Table 2a and in columns 1 to 4 of Table 2b have expected positive signs: dissimilarity in terms of social class composition, education or age increases dissimilarity in terms of preferences for redistribution, general trust and corruption significantly. Dissimilarity in terms of occupation however has an unexpected negative and significant sign of the Manhattan distance of corruption between pairs of PSUs. Similarly, religious heterogeneity actually contributes negatively to the dissimilarity of the general trust index between pairs of PSUs. All excluded covariates in columns 5 to 8 of Table 2b and in Table 2c also have expected signs: dissimilarities between PSU pairs in terms of age, education or religion increases PSUs' dissimilarities in terms of occupation; and in female labor force participation.

We can see from Table 2a and 2b that geography and history both play a role in cultural dissimilarity. Physical distance increases cultural dissimilarities, while belonging to the same country or former empire or sharing a border tones them down. This is true of all values under investigation, namely preferences for redistribution and general trust, but also of economic outcomes, such as the prevalence of corruption or female labor force participation. The effect of physical distance, which in our theoretical model is a proxy for local social interactions, remains

important, even once past and modern geopolitical considerations are taken into account, except in the case of general trust and occupation. The effect of past empires has a consistent and robust effect across all of the independent variables under consideration. This is also true of today's geopolitical borders. The effect is quite strong. It is useful to express the impact of each variable in terms of geographical distance. For example, belonging to the same country reduces our measure of dissimilarity in terms of general trust by the equivalent of 995 km<sup>6</sup>. Belonging to 2 different but contiguous countries has a comparable negative effect on this measure: it decreases it by the equivalent of 663 km. Interestingly, the effect of past history is even stronger than the impact of today's geopolitical borders: having been integrated in the same former empire during more than 100 years decreases the measure of general trust by the equivalent of 1106 km. For female labor force participation, the strongest effect comes from belonging to the same country today: it decreases the dissimilarity between locations pair by an equivalent of nearly 3100 km. Belonging to two adjacent countries has a rather strong effect as well, by reducing the dissimilarity between PSU pairs in terms of female labor force participation by the equivalent of 2322 km, while the effect of former integration in empires is to reduce it by "only" slightly more than 1000 km.

#### ***4.2. More recent history: the effect of the USSR and Yugoslavia***

We add to the specification in (9) two dummy variables investigating further the role of past, albeit more recent history: USSR and Yugoslavia, which take value 1 if both PSUs in the pair used to belong to, respectively, the former USSR and former Yugoslavia. Results are displayed in Table 3. Results are very clear-cut in the case of the former USSR: dissimilarity in all of the dimensions measured in this paper is larger within the former USSR than it is on average when the whole sample is considered. Results are more contrasted in the case of the former Yugoslavia. PSUs in countries of the former Yugoslavia are more similar to each other in terms of preferences for redistribution and corruption, while the opposite is true for all our other indicators. In any case, the contrast with the effect of the variable Same Empire is flagrant. Locations that belonged to the former Russian, Ottoman, Habsburg or Prussian empires for more than a 100 years are more similar to each other than locations that were integrated in the former

---

<sup>6</sup> Computed as  $23.22 * 0.009 / (0.021 / 100)$ , 23.22 is equivalent to a 1 % change in the average distance.

USSR or Yugoslavia, despite the fact that these last two historical episodes are closer, albeit briefer, events in the countries' past.

**Table 3: The Effect of the former USSR and Yugoslavia**

		1	2	3	4	5	6
	Pref. redistribution	General trust	Corruption	Occupation	Proportion active women	Proportion women active in 1989 and still	
Distance	-0.006*** [0.001]	0.022*** [0.001]	0.001 [0.001]	0.010*** [0.001]	0.010*** [0.000]	0.007*** [0.000]	
Same Country	-0.084*** [0.003]	-0.015*** [0.003]	-0.057*** [0.003]	-0.029*** [0.002]	-0.014*** [0.001]	-0.010*** [0.001]	
Contiguous	-0.050*** [0.002]	-0.009*** [0.001]	-0.001 [0.002]	-0.026*** [0.001]	-0.010*** [0.001]	-0.003*** [0.001]	
Same Empire	-0.030*** [0.002]	-0.009*** [0.001]	-0.017*** [0.002]	-0.010*** [0.001]	-0.004*** [0.001]	-0.001* [0.001]	
USSR	0.034*** [0.003]	0.005* [0.003]	0.022*** [0.003]	0.003 [0.002]	0.007*** [0.001]	0.012*** [0.001]	
Yugoslavia	-0.145*** [0.003]	0.046*** [0.003]	-0.016*** [0.003]	0.022*** [0.002]	0.013*** [0.001]	0.002* [0.001]	
Observations	963966	963966	963966	963966	957036	957036	
R-squared	0.089	0.07	0.101	0.232	0.045	0.034	

Notes to Table 3:

All dependent variables are Manhattan Distance measures of dissimilarity between PSU pairs. Distance is expressed as the logarithm of the physical distance between PSUs in km. Regressions columns 1 to 4 control for dissimilarity between pairs of PSUs in terms of religious affiliation, occupation, social class composition, education, age, as well as country fixed effects. Regressions in columns 5 to control for dissimilarity between pairs of PSUs in terms of religious affiliation, education, age, as well as country fixed effects.

Same Empire takes value 1 if both PSUs have belonged to the same Empire (as defined in section 4) for more than a 100 years. USSR is a dummy variable that takes value 1 if both PSUs in the pair used to belong to the former USSR. Yugoslavia is a dummy variable that takes value 1 if both PSUs in the pair used to belong to the former Yugoslavia.

The former USSR comprises Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, the Russian Federation, Tajikistan, Ukraine and Uzbekistan. Former Yugoslavia consists of Bosnia and Herzegovina, Croatia, FYR Macedonia, Montenegro, Serbia and Slovenia.

\*\*\* indicates significance at 1%, \* significance at 10%. Standard errors are in brackets.

## 5. Extensions

### 5.1. The effects of EU integration

This cultural gravity model can be used in order to investigate the effect of European Union integration. Many countries in our sample are either member states of the European Union or candidate countries. The Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, the Slovak Republic and Slovenia are member states of the EU since 2004; Bulgaria and Rumania since 2007. Other countries, such as Croatia, the Former Yugoslav Republic of Macedonia and Turkey are candidate countries. We construct several dummy variables to reflect the EU status of each country in our sample, and we evaluate the impact of EU status on cultural dissimilarity and our measures of economic outcomes heterogeneity. We construct two dummy variables, which take value one if the countries to which the PSU pair belongs are both in the EU (EU), or are both candidate states (EU candidates). We the construct a dummy variable that takes value one if one PSU belongs to a member states and the other to a candidate country (EU-candidate EU). The last dummy variable (EU- nonEUocand) takes value one if one PSU belongs to a EU member states and the other is neither a member states nor a candidate. The purpose of these last two variables is hence to reflect whether countries that are not yet in the EU but are candidates are culturally and economically more similar to EU member states rather than countries which do not have any perspective of adhesion yet. An interesting implication is the nature of integration in a large cultural ensemble.

**Table 4: The Effects of actual and future EU integration. Results of the Gravity Model**

	1 Pref. redistribution.	2 Corruption	3 Gen. trust	4 Occupation	5 Prop. active women	6 Prop. women active in 1989 still active
Distance	-0.003*** [0.001]	0.015*** [0.001]	0.000 [0.001]	0.005*** [0.001]	0.009*** [0.000]	0.005*** [0.000]
Same country	-0.101*** [0.003]	-0.001 [0.003]	-0.063*** [0.003]	-0.020*** [0.002]	-0.011*** [0.001]	-0.006*** [0.001]
Contiguous	-0.065*** [0.002]	-0.006*** [0.001]	-0.002 [0.002]	-0.027*** [0.001]	-0.010*** [0.001]	-0.003*** [0.001]
<b>EU</b>	<b>0.095***</b>	<b>-0.168***</b>	<b>-0.005</b>	<b>-0.036***</b>	<b>-0.015***</b>	<b>-0.017***</b>

	[0.010]	[0.008]	[0.010]	[0.006]	[0.004]	[0.003]
EU candidates	-0.071***	-0.003	-0.022***	0.013***	-0.007***	0.000
	[0.006]	[0.005]	[0.006]	[0.004]	[0.002]	[0.002]
EU-candidate EU	0.116***	-0.016***	-0.025***	0.029***	0.010***	0.011***
	[0.004]	[0.003]	[0.004]	[0.003]	[0.002]	[0.001]
EU- nonEUocand	0.057***	-0.039***	0.011**	0.010***	-0.003*	0.003**
	[0.005]	[0.004]	[0.005]	[0.003]	[0.002]	[0.002]
Observations	963966	963966	963966	963966	957036	957036
R-squared	0.089	0.073	0.101	0.233	0.045	0.036

Notes to Table 4:

All dependent variables are Manhattan Distance measures of dissimilarity between PSU pairs. Distance is expressed as the logarithm of the physical distance between PSUs in km. All regressions control for country fixed effects. Regressions in columns 1 to 3 control for dissimilarity between pairs of PSUs in terms of religious affiliation, education, age, occupation and social class composition. Regressions in columns 4 to 6 control for dissimilarity between pairs of PSUs in terms of religious affiliation, education and age. \*\*\* indicates significance at 1%, \* significance at 10%. Standard errors are in brackets.

The negative coefficient on the EU dummy variable indicates a strong convergence between EU member states in terms of our economic outcomes indicators: occupation, female labor force participation and corruption. The effect is quite large. For example, belonging to the EU reduces the dissimilarity of economic occupations between two PSUs by the equivalent of 1672 km. For the proportion of women of working age that are actually active, the equivalent number is 18067. On the contrary, EU members seem to diverge from non EU members in terms of occupation, corruption and the second indicator of female labor force participation. These results seem to indicate that EU membership has a centrifuge effects for its members and a centripetal effects on countries that are neither member states nor candidates on economic outcomes. Recall that these results control for distance and other geographical variables such as contiguity or same country dummy. Results in what regards candidate countries are less clear. Compared with countries that are neither in the EU nor candidates, EU candidate countries are less dissimilar to EU countries in terms of corruption and female activity, but results concerning the other variables are not very dissimilar.

Our data does not allow us to identify whether our results is due to a convergence between EU countries themselves, and to some extent with candidate countries or a selection effect, but at least it indicates that there is something more than geographical proximity in the EU. This leads to the following question: Is Turkey in Europe?

## 5.2. Is Turkey in Europe?

Is Turkey closer to European member states in terms of the values and economic outcomes we measure in this paper, or is it closer to other non European countries of our sample, including the other candidate to the EU?

We keep only PSU pairs in which at least one PSU is situated in Turkey. Results are displayed in Table 5a.

**Table 5a. Is Turkey in Europe? Results of the cultural gravity model**

	1	2	3	4	5	6
	Pref. redistribution	Corruption	Gen. trust	Occupation	Prop. active women	Prop. women active in 1989 still active
Distance	0.017*** [0.004]	-0.016*** [0.005]	-0.041*** [0.004]	0.076*** [0.003]	0.040*** [0.002]	0.019*** [0.001]
MD religion	-0.042*** [0.005]	0.003 [0.005]	0.016*** [0.004]	0.006* [0.003]	-0.010*** [0.002]	0.002 [0.002]
MD occupation	0.043*** [0.006]	-0.015*** [0.006]	0.004 [0.005]			
MD social class	0.014*** [0.005]	0.032*** [0.005]	0.051*** [0.004]			
MD education	0.061*** [0.005]	0.036*** [0.005]	0.003 [0.004]	0.339*** [0.003]	0.039*** [0.002]	0.049*** [0.002]
MD age	-0.003*** [0.000]	-0.003*** [0.000]	-0.001*** [0.000]	0.008*** [0.000]	0.000*** [0.000]	0.001*** [0.000]
Same country	0 [0.000]	0 [0.000]	0 [0.000]	0 [0.000]	0.054*** [0.009]	0.01 [0.007]
EUNonEU	0.259*** [0.018]	0.207*** [0.019]	0.104*** [0.016]	-0.058*** [0.013]	0.118*** [0.006]	0.043*** [0.005]
EU candidates	-0.136*** [0.013]	-0.257*** [0.013]	-0.068*** [0.012]	-0.069*** [0.009]	0.007 [0.006]	-0.025*** [0.005]
Observations	68175	68175	68175	68175	67925	67925
R-squared	0.222	0.121	0.035	0.216	0.082	0.091

Notes to Table 5a:

All dependent variables are Manhattan Distance measures of dissimilarity between PSU pairs. Distance is expressed as the logarithm of the physical distance between PSUs in km. Regressions control for country fixed effects. \*\*\* indicates significance at 1%, \* significance at 10%. Standard errors are in brackets

Results indicate that Turkey is more dissimilar to the new EU member states than to the rest of the sample in almost every respect, except preferences for redistribution. However, Turkey is

very similar to other EU candidate states. This result is certainly due to the fact that candidate countries include countries that were part of the Ottoman Empire. It may be distinguishing the first wave of new EU members, and the second wave, which includes Bulgaria and Romania, two countries that belonged (partly for the case of Romania) to the former Ottoman Empire. Results are displayed in Table 5b. The same result holds: Turkey is more dissimilar to new EU members than to the rest of the sample, even when one distinguishes new EU members that were former satellites of the Ottoman Empire. Nevertheless, it remains true that Turkey is quite similar to other EU candidates.

**Table 5b: Is turkey in the (very new) Europe?**

	1	2	3	4	5	6
	Pref. redistribution	Corruption	Gen. trust	Occupation	Prop. active women	Prop. women active in 1989 still active
Distance	0.017*** [0.004]	-0.016*** [0.005]	-0.041*** [0.004]	0.076*** [0.003]	0.040*** [0.002]	0.019*** [0.001]
MD religion	-0.042*** [0.005]	0.003 [0.005]	0.016*** [0.004]	0.006* [0.003]	-0.010*** [0.002]	0.002 [0.002]
MD occupation	0.043*** [0.006]	-0.015*** [0.006]	0.004 [0.005]			
MD social class	0.014*** [0.005]	0.032*** [0.005]	0.051*** [0.004]			
MD education	0.061*** [0.005]	0.036*** [0.005]	0.003 [0.004]	0.339*** [0.003]	0.039*** [0.002]	0.049*** [0.002]
MD age	-0.003*** [0.000]	-0.003*** [0.000]	-0.001*** [0.000]	0.008*** [0.000]	0.000*** [0.000]	0.001*** [0.000]
Same country	0 [0.000]	0 [0.000]	0 [0.000]	0 [0.000]	0.047*** [0.008]	0.054*** [0.007]
FirstWaveEU	0.259*** [0.018]	0.207*** [0.019]	0.104*** [0.016]	-0.058*** [0.013]	0.064*** [0.006]	0.033*** [0.005]
SecondWaveEU	0.140*** [0.014]	-0.381*** [0.014]	0.001 [0.012]	0.052*** [0.010]	0.118*** [0.006]	0.043*** [0.005]
EU candidates	-0.136*** [0.013]	-0.257*** [0.013]	-0.068*** [0.012]	-0.069*** [0.009]	0.007 [0.006]	-0.025*** [0.005]
Observations	68175	68175	68175	68175	67925	67925
R-squared	0.222	0.121	0.035	0.216	0.082	0.091

*Notes to Table 5b:*

*All dependent variables are Manhattan Distance measures of dissimilarity between PSU pairs. Distance is expressed as the logarithm of the physical distance between PSUs in km. Regressions control for country fixed effects. \*\*\* indicates significance at 1%, \* significance at 10%. Standard errors are in brackets*

### ***5.3. Internet and cultural (and economic) homogenization***

Estimating the impact of internet access on cultural values, as well as on economic activity, runs into an endogeneity problem. People who chose and can afford to have internet may be very different culturally, as well as in their patterns of economic activity, from people who do not. A candidate instrumental variable for internet access would be whether technologies for internet access have been installed in the region. We construct a dummy equal to one if at least one household in the PSU has internet access at home (instead of using the mean value of internet users in the PSU). This variable is intended to reflect whether internet access is available in the village or town in which each household lives. Less than half of the PSU pairs both have internet access in our sample. Of course, the roll out of installing internet technologies is not exogenous either. Our dummy variables does not alleviate this issue, but alleviates the endogeneity of internet access at the household level.

In the table that follows, the variable Internet indicates whether cultural dissimilarity and the heterogeneity of economic outcomes in terms of corruption, occupation, or female labor force participation is lessened when both PSU have internet access. In other words, it is a measure of cultural and economic homogenization induced by internet access. Our results indicate that internet does indeed have a strong homogenizing effect. Although the coefficient might be upward biased because of the endogeneity problem mentioned earlier, the effect is quite strong. For example, having internet access in both PSUs is equivalent to decreasing the physical distance by 2786 km in terms of dissimilarity between PSU in their general trust attitude and by more than 5280 km in terms of their heterogeneity in their occupation structure.

**Table 6a: Internet and cultural (and economic) homogenization. Results of the gravity model**

		1	2	3	4	5	6
	Pref. redistribution	Corruption	Gen. trust	Occupation	Prop. active women	Prop. women active in 1989 still active	
Distance	-0.001 [0.001]	0.001* [0.001]	0.020*** [0.001]	0.004*** [0.001]	0.007*** [0.000]	0.006*** [0.000]	
MD religion	0.007*** [0.001]	-0.006*** [0.001]	0.015*** [0.001]	0.038*** [0.000]	0.010*** [0.000]	0.005*** [0.000]	
MD occupation	0.062*** [0.002]	0.059*** [0.002]	-0.023*** [0.001]				
MD social class	0.025*** [0.001]	0.029*** [0.001]	0.028*** [0.001]				
MD education	0.034*** [0.001]	0.042*** [0.001]	0.031*** [0.001]	0.183*** [0.001]	0.047*** [0.001]	0.013*** [0.000]	
Abs diff age	0.001*** [0.000]	0.002*** [0.000]	0.003*** [0.000]	0.014*** [0.000]	0.000*** [0.000]	0.002*** [0.000]	
Same country	-0.109*** [0.003]	-0.063*** [0.003]	-0.013*** [0.003]	-0.036*** [0.002]	-0.016*** [0.001]	-0.011*** [0.001]	
Contiguous	-0.060*** [0.002]	-0.003* [0.002]	-0.009*** [0.001]	-0.029*** [0.001]	-0.010*** [0.001]	-0.003*** [0.001]	
Internet dummy	-0.010*** [0.001]	0.000 [0.001]	-0.024*** [0.001]	-0.091*** [0.001]	-0.034*** [0.000]	-0.014*** [0.000]	
Observations		963966	963966	963966	963966	957036	957036
R-squared		0.086	0.101	0.07	0.245	0.051	0.036

Notes to Table 6a:

All dependent variables are Manhattan Distance measures of dissimilarity between PSU pairs. Distance is expressed as the logarithm of the physical distance between PSUs in km. Regressions control for country fixed effects. \*\*\* indicates significance at 1%, \* significance at 10%. Standard errors are in brackets

One interesting question to study is to which sort of values internet makes users converge to. Using a Manhattan Distance measure only gives an idea of the scope of cultural differences, not in which direction the values evolve. An important question is whether internet use make people converge towards “universal” values, or towards communitarianism. In order to get an idea, I re run the above regressions by including an interaction between the Same country dummy and the internet dummy. This interaction term gives an indication as to whether internet users converge more towards each other within the same country or whether this effect is oblivious of political borders. Results are displayed in Table 6b. The coefficient on the interaction is negative and significant in columns 1 to 3, indicating that internet reduces cultural distance, that is to say our

measures of dissimilarity in general trust, preferences for redistribution, as well as corruption, all the more so *within* the same country, but the opposite is true for the heterogeneity of economic outcomes in terms of occupation and female labor force participation.

**Table 6b: Does internet homogenize within or across borders?**

	1 Pref. redistribution	2 Corruption	3 Gen. trust	4 Occupation	5 Prop. active women	6 Prop. women active in 1989 still active
Distance	-0.001 [0.001]	0.002* [0.001]	0.020*** [0.001]	0.004*** [0.001]	0.007*** [0.000]	0.006*** [0.000]
MD religion	0.008*** [0.001]	-0.006*** [0.001]	0.015*** [0.001]	0.038*** [0.000]	0.010*** [0.000]	0.005*** [0.000]
MD occupation	0.062*** [0.002]	0.060*** [0.002]	-0.023*** [0.001]			
MD social class	0.025*** [0.001]	0.029*** [0.001]	0.028*** [0.001]			
MD education	0.034*** [0.001]	0.042*** [0.001]	0.031*** [0.001]	0.183*** [0.001]	0.047*** [0.001]	0.013*** [0.000]
Abs diff age	0.001*** [0.000]	0.002*** [0.000]	0.003*** [0.000]	0.014*** [0.000]	0.000*** [0.000]	0.002*** [0.000]
Same country	-0.096*** [0.004]	-0.033*** [0.004]	0.001 [0.004]	-0.068*** [0.003]	-0.023*** [0.002]	-0.013*** [0.001]
Contiguous	-0.060*** [0.002]	-0.003 [0.002]	-0.008*** [0.001]	-0.029*** [0.001]	-0.010*** [0.001]	-0.003*** [0.001]
Internet dummy	-0.010*** [0.001]	0.001 [0.001]	-0.023*** [0.001]	-0.093*** [0.001]	-0.034*** [0.000]	-0.014*** [0.000]
Internet*country	-0.024*** [0.005]	-0.052*** [0.005]	-0.026*** [0.004]	0.056*** [0.003]	0.011*** [0.002]	0.005*** [0.002]
Observations	963966	963966	963966	963966	957036	957036
R-squared	0.086	0.101	0.07	0.245	0.051	0.036

Notes to Table 6b:

All dependent variables are Manhattan Distance measures of dissimilarity between PSU pairs. Distance is expressed as the logarithm of the physical distance between PSUs in km. Regressions control for country fixed effects. \*\*\* indicates significance at 1%, \* significance at 10%. Standard errors are in brackets

## 6. Conclusion

## 7. References

- Ivan T. Berend. *History Derailed: Central and Eastern Europe in the Long Nineteenth Century*. Berkeley: University of California Press, 2003. xix, 330 pp
- Botticini, Maristella & Eckstein, Zvi, 2006a. "From Farmers to Merchants, Voluntary Conversions and Diaspora: A Human Capital Interpretation of Jewish History," CEPR Discussion Papers 5571, C.E.P.R. Discussion Papers.
- Botticini, Maristella & Eckstein, Zvi, 2006b. "Path Dependence and Occupations, Boston University - Department of Economics - The Institute for Economic Development Working Papers Series DP-154, Boston University - Department of Economics.
- Bruegger, Beatrix, Rafael Lalive and Josef Zweimuller (2008\_) Does culture affect unemployment? Evidence from the Barriere des Roestis
- de Paula, Auro & Jos A. Scheinkman, 2007. "The Informal Sector," NBER Working Papers 13486, National Bureau of Economic Research, Inc.
- Durlauf, Steven N, 1996. " A Theory of Persistent Income Inequality," *Journal of Economic Growth*, Springer, vol. 1(1), pages 75-93, March
- Guiso, Luigi, Paola Sapienza & Luigi Zingales, 2008. "Long Term Persistence," NBER Working Papers 14278, National Bureau of Economic Research, Inc.
- Head, Thierry and Keith Head (2008). Detection of local interactions from the spatial pattern of names in France. *Journal of Regional Science*, vol 48, number, pp 67-95.
- Kelejina, Harry H., Peter Murrell and Oleksandr Shepotyloe (2007). *Spatial Spillovers in the Development of Institutions*.
- Lindbeck, Assar & Sten Nyberg & Jörgen W. Weibull, 2003. "Social Norms and Welfare State Dynamics," *Journal of the European Economic Association*, MIT Press, vol. 1(2-3), pages 533-542, 04/05 "Incentives and Social Norms in Household Behavior," *American Economic Review*, American Economic Association, vol. 87(2), pages 370-77, May.
- Manski, Charles F. (1993), "Identification of Endogenous Social Effects: The Reflection Problem" *The Review of Economic Studies*, 60(3), 531–542
- Bertrand, Marianne, Erzo F. P. Luttmer & Sendhil Mullainathan, 2000. "Network Effects And Welfare Cultures," *The Quarterly Journal of Economics*, MIT Press, vol. 115(3), pages 1019-1055, August

Glaeser, Edward L & Sacerdote, Bruce & Scheinkman, Jose A, 1996. "Crime and Social Interactions," The Quarterly Journal of Economics, MIT Press, vol. 111(2), pages 507-48, May

Tabellini: Institutions and culture, 2007

tabellini scope of cooperation, 2008, QJE, forthcoming

which itself stems fro and from the multi-ethnic nature of many countries in the sample.

## 8. Appendix

**Table A1: Geographical and temporal delimitations of Empires in our sample**

	<b>600 years or more</b>	<b>300 to 500 years</b>	<b>200 years or less</b>
<b>Ottoman Empire</b>	Bulgaria (except part which became independent in 1878), FYROM, Turkey	Albania, Bosnia, Bulgaria which became independent in 1878, Bessarabia, Crimea, Moldavia, outer Montenegro, Serbia except Vojvodina, Wallachia	Transylvania, Vojvodina, southern Hungary, eastern Croatia
<b>Habsburg Empire</b>	Slovakia, Hungary (except southern-ottoman Hungary)	Croatia (except Dalmatia), Czech Republic, Southern (ottoman) Hungary , Polish Silesia, Slovenia, Transylvania, Vojvodina	Dalmatia, Galicia, Habsburg Poland (Krakow, Rzeszow), Bosnia, western Ukraine (Lviv)
<b>Prussia</b>	Pomerania (Poland)	Estonia, Latvia, Polish Silesia, Royal Prussia	Polish Silesia, Kaliningrad, Klaipeda (Lithuania)
<b>Russian Empire</b>		Russia (except Kaliningrad), Ukraine (except Crimea and Kouban)	Armenia, Azerbaijan, Belarus, Baltic states (except Klaipeda), Georgia, , Kazakhstan, Kyrgyzstan, Moldova, eastern-central Poland (Warsaw, Lodz), rest of Ukraine, Tajikistan

**Table A2: Descriptive statistics**

Variable	Mean	Std. Dev.	Min	Max
<b>Dependent variables</b>				
dist	2321.998	1892.778	0	14397.8
ln (distance)	7.340031	1.035817	0	9.5749
MD preferences redistribution	0.6187724	0.4451283	0	2
MD corruption	0.9744148	0.4848529	0	2
MD general trust	0.8738906	0.3855833	0	2
MD occupation	0.9366432	0.3202894	0.0374922	2.477694
Abs. diff Proportion active women	0.2479292	0.1806491	0	1
Abs. diff Proportion women active in 1989 still active	0.1885439	0.1428231	0	0.9411765
<b>Independent variables</b>				
	Mean	Std. Dev.	Min	Max
MD religion	0.9368525	0.7540388	0	2
MD social class	0.6488456	0.3626464	0	2
MD education	0.896283	0.3744252	0.0049547	2
(abs) difference age	8.113819	6.067242	0.0000114	49.45921
Same country	0.0350334	0.1838643	0	1
contiguous	0.1300554	0.3363646	0	1
Same Empire	0.1495613	0.3566411	0	1
USSR	0.2465906	0.4310266	0	1
Yugoslavia	0.0465265	0.2106226	0	1
EU	0.1288956	0.3350846	0	1
candEU	0.0077804	0.0878626	0	1
EUCandEU	0.077648	0.267617	0	1
EUnonEUoc~d	0.4088422	0.4916203	0	1
Dinternet	0.4328379	0.495469	0	1