Genetic Map for Economists[1]

FIRST DRAFT

April 17, 2006

Paola Giuliano
Research Department, IMF; IZA

Antonio Spilimbergo
Research Department, IMF; CEPR; WDI

Giovanni Tonon
Dana-Farber Cancer Institute, Harvard University

Abstract

What does genetic distance between populations measure? Is genetic distance a good proxy for culture and a valid instrument to disentangle the causal relationship between culture and economic outcomes? Taking an interdisciplinary approach, we examine how economists may interpret the correlation between genetic distance, and cultural and economic variables. We argue that currently used measures of genetic distance are a poor proxy for cultural differences. Rather, genetic distance, being determined, among other things, by geographical barriers, reflects transport costs between countries. To demonstrate this point, we construct a new measure of geographic distance within Europe that takes into account the existence of major geographical barriers. We show that this measure explains both genetic distance and trade between European countries.

JEL Classification Numbers: Z10, F10
Keywords: transport costs, instrumental variables, genetics, trade.
Authors' E-mail address: pgiuliano@imf.org; aspilimbergo@imf.org; Giovanni_Tonon@dfci.harvard.edu.

[1] The views expressed in this paper are those of the authors and do not necessarily represent those of the IMF, its Board of Directors, or the views of the Dana Farber Cancer Institute. We thank Andrea Bianchi, Oya Celasun, Simon Johnson, Paolo Mauro, Arvind Subramanian, Antonio Terraciano and Alessandro Vanzetti for useful discussions, and Zhaogang Qiao and José Romero for excellent research assistantship.

I. INTRODUCTION

Cultural factors have a strong influence on economic and social phenomena. This proposition has a long intellectual history in Western thought going back to Greek and Roman literatures. More recently, social scientists have argued that cultural innovations have led to the development of capitalism (Weber, 1958) or that different historical and culture experiences in Italian regions have led to different development paths (Putnam, 1993). Despite this long tradition, the analysis of causal relationships between culture and economic outcomes has proven challenging because of the several interrelations between culture and economics.

Trying to disentangle the causal relationship between culture and economic factors, social scientists have used two research strategies. The first strategy (*deductive*) consists in analyzing the content of a particular culture and explaining why specific cultural factors (for instance motivation to work, saving rates, respect for the laws) have led to different economic outcomes. For instance, Weber (1958) dissected the logic behind the early protestant morality and showed how the protestant ethic favored the development of a `rational mentality,' which would be a precondition for capitalism. The advantage of this strategy is that specific mechanisms of transmission are identified and every step can be brought to test.[2] The second strategy (*inductive*) consists in finding (conditional) correlations between socio-economic outcomes and culture indicators. This second strategy, which is presently the standard in economic literature, has the advantage of avoiding ad-hoc speculations on how culture has an impact on economic phenomena but leaves open the question of reverse causality.

To give a causal interpretation to the correlation between culture variables and economic outcomes, economists have relied on instrumental variables, which are supposedly correlated with cultural variations but without direct impact on economic phenomena. Finding the right instruments remains the biggest challenge in the analysis of culture. Tabellini (2005) has used

[2] Note that Max Weber in his "The Economic Ethic of the World Religions" (1958) extends his analysis to Confucian, Hinduist, Buddhist, Christian, and Islamist religions. In each religion, Weber is interested in the `economic ethic,' i.e. "the practical impulses for action which are found in the psychological and pragmatic contexts of religions." In particular, he derives from the ethic of each religion the consequence in terms of individual motivations and social stratification, which, in turn, determine economic outcomes. Following this approach, Weber argued that capitalist development was unlikely to occur in many Asian countries(!).

literacy rates at the end of the 19th century and political institutions in place over the past several centuries; Putnam (1993) has used differences prevailing in local governments in Italian regions several centuries ago. Two recent papers propose genetic distance as an instrument or as a proxy for culture (Guiso, Sapienza and Zingales, 2005 and Spolaore and Wacziarg, 2006).

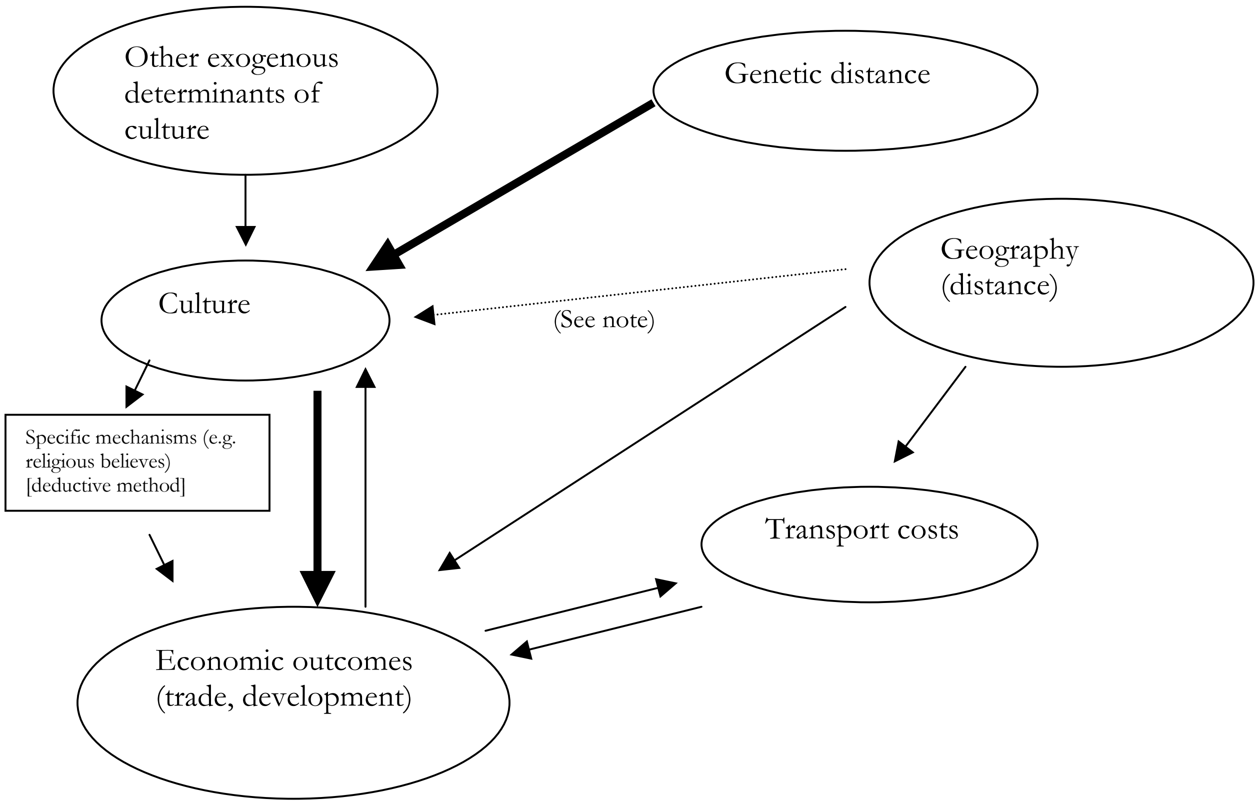
To quantify the degree of genetic differentiation among two or more populations of the same specie, geneticists often use an index defined genetic distance.[3] In economics, it has been argued that this measure is an ideal proxy for culture (Spolaore and Wacziarg, 2006) because genetically similar populations are more likely to share similar languages, habits and norms. Moreover, being related to Neolithic migrations, genetic distance would seem unlikely to be related to current economic activity (Guiso, Sapienza and Zingales, 2005).

The use of instrumental variables is not simply a robustness test to control for potential endogeneity of some regressors or for measurement errors in variables but it is the only way economists can interpret a (partial) correlation as a causal relationship. In this context, genetic distance may seem the ideal instrument for cultural variables because: 1) prima facie, genetic distance seems unlikely to have a direct effect on economic variables (exclusion condition); 2) genetic distance seems correlated with cultural variables (relevance of the instrument); finally, 3) there appears to exist no third variables determining both economic variables and genetic distance (no omitted third variables). The following chart summarizes these assumptions: [4]

---

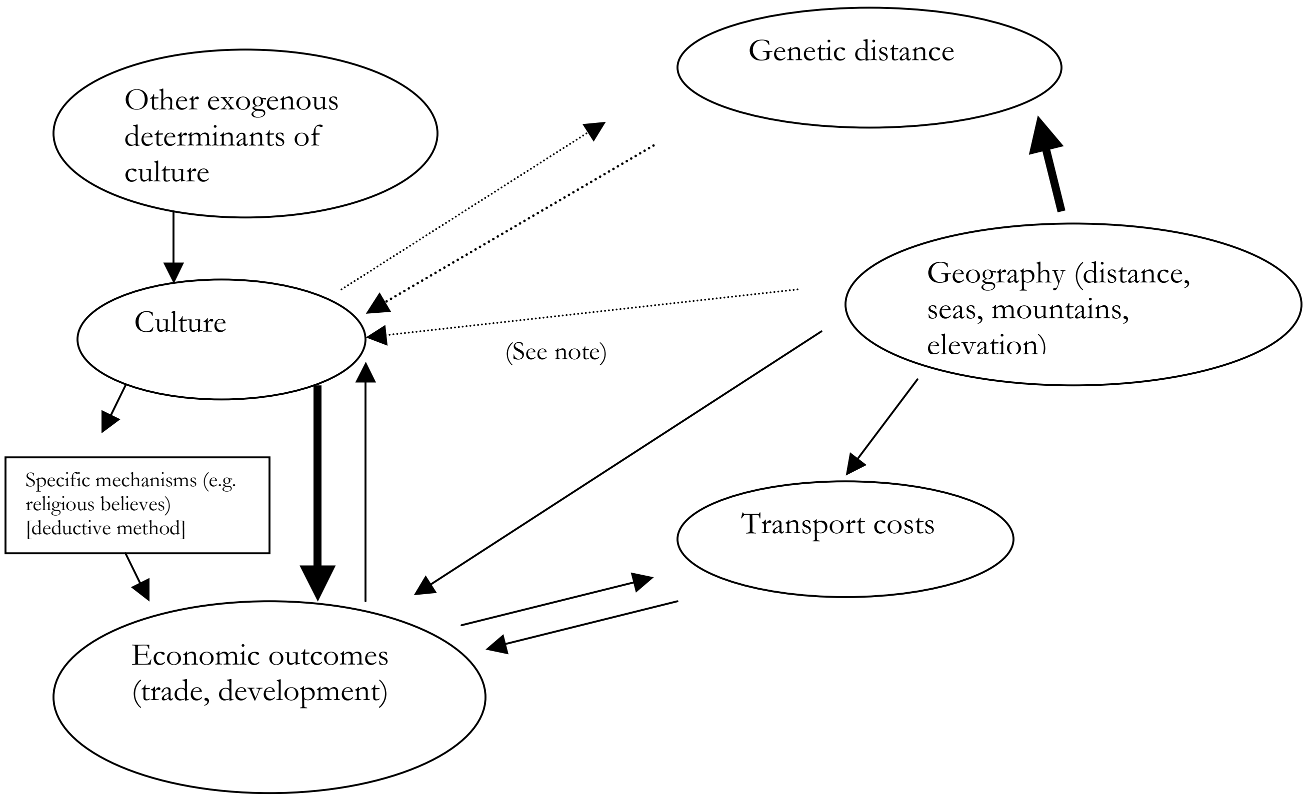
[3] Section 2 below defines the measure of genetic distance.

[4]For simplicity, the chart above ignores institutions and their link with economic outcomes, which has been also extensively studied. Note also that in antiquity there was a widespread belief that geography had a strong direct impact on cultural factors (dashed line in the chart). Consider the following quote from Hippocrates' "On airs, waters and places," chapter 12: "I say then that Asia differs very much from Europe as to the nature of all things, both with regard to the productions of the earth and the inhabitants, for everything is produced much more beautiful and large in Asia; the country is milder, and the dispositions of the inhabitants also are more gentle and affectionate. The cause of this is the temperature of the seasons, because it lies in the middle of the risings of the sun towards the east, and removed from the cold (and heat), for nothing tends to growth and mildness so much as when the climate has no predominant quality, but a general equality of temperature prevails." These chapters contain a detailed comparison between Asia and Europe and some considerations on their inhabitants. As the treatise dates back to the fifth century B.C., this may be one of the most ancient comparisons between continents.



This paper questions the validity of the last two assumptions (that there are no omitted variables, and that genetic distance and cultural differences are strongly correlated) and proposes a different interpretation of the correlation between genetic distance and economic activity, specifically trade. First, genetic distance and economic outcomes are both influenced by geographical variables. Present genetic structure in Europe was shaped by natural selection, migrations and genetic drift, the latter ones largely determined by geographical impediments. Similarly, after several millennia and despite advancements in transportation, the mountains, the rivers, and the seas that shaped past migrations and genetic drift continue to have an impact on modern transport costs and ultimately on trade. Second, the hypothesis that genetic and cultural distances are strongly correlated is not consistent with the evidence from anthropological studies, genetic analysis, and simple correlations provided below.

In conclusion, genetic distance between two nations is correlated with modern transport costs while it is only loosely correlated with cultural distances; hence, the use of genetic distance as an instrument for cultural distance or as a proxy for culture is not warranted. This is not to say that culture does not play an important role in economic phenomena, but that the use of genetic distance as an instrument or as a proxy for culture is inappropriate. The following chart illustrates the points further:



We make our case by presenting evidence from other disciplines, including genetics and social psychology, and by studying the correlation between genetic distance and trade within European countries, because the richness of data for genetic distance and on trade allows a more detailed analysis. In particular, we address the following questions: (1) what do the available measures of genetic distance really measure? (2) Why does genetic distance explain trade flows so well? (3) If culture and genetic distance are only loosely correlated, what does genetic distance proxy for? (4) Are there any other uses for genetic distance in economics?

The rest of the paper is organized as follows. Section II provides an overview of the available measures of genetic distance, highlighting how they are calculated, what they really measure and their relationship with physical anthropology data, including anthropometric characters like stature or qualitative traits such as eye color or skin pigmentation. Section III shows that genetic distance may explain very well trade between European countries in a standard gravity equation with (log)-distance as a proxy of transport costs, but becomes insignificant once transport costs or other variables capturing geographical features are introduced. Section IV shows that commonly used variables in social psychology on cultural differences, including McCrae and Hofstede indices, are poorly correlated with genetic distance. Section V discusses alternative uses of genetic distance in economics; section VI concludes.

II. WHAT IS IN THE GENES?

Population genetics studies populations' genetic composition and their changes over time.[5] In its simplest form, the fundamental measurement in population genetics is the frequency at which alleles are found at any specific *gene locus* (allele frequency).[6] Population genetics studies genes that are present in at least two different forms (alleles) in the population.

Interestingly, although not all alleles occur in all human populations, differences in alleles within local human populations are much greater than among different populations. Specifically, 93% of total human variability is found within local populations. The remaining 7% is found between populations (Rosenberg et al., 2002). As noted by Lewontin, "if everyone on earth

---

[5] Although the biological concept and the political concept of "populations" are often used as synonymous, they are clearly different. Even in Europe, where the two concepts often overlap, there are noticeable exceptions: for instance, the genetic distance between a Sardinian and the rest of Italy (221) is more than twice than the difference between Italy and Norway (88) or Italy and Sweden (95).

[6] With the term *gene locus* (or, for simplicity, gene) we intend a sequence of DNA that encodes for a protein, and with the term *allele* we consider a particular form of a specific gene. Often alleles are distinguished for their effects on the phenotype (e.g., morphological, physiological or biochemical characteristics of an individual – or group of like individuals – that differ in this respect from other individuals), or, simply, for differences in pair sequence. Several different methods have been used to measure the genetic composition of a population. Some of these techniques are directly linked to DNA alterations, as for example restriction fragment length polymorphism (RFLPs), single nucleotide polymorphisms (SNPs) and mutations, micro satellites and most recently haplotypes. Classical analysis instead measured the result of DNA alterations, that is, protein variation. The most extensive and comprehensive studies on variants have been performed on protein polymorphism. The ongoing Human Genome Diversity Project (HGDP) and the International HapMap Project will soon provide a wealth of data and information linked directly to the DNA status, but the results available so far and the analysis performed on these data are not exhaustive. Preliminary analysis, however, supports the notion that the major tenets of the classical protein polymorphism analysis, as presented in its more comprehensive form by Cavalli-Sforza et al. in 1994, correlates closely with this new, more extensive scrutiny. For this reason, in the present study, we have relied on the data as provided in Cavalli-Sforza.

becomes extinct except for the Kikuyu of East Africa, about 90% of all human variability would still be present in the reconstituted species" (Lewontin, 1984).

Subsets of this specific group of genes that varies between populations are used to reconstruct the evolutionary history of populations. Moreover, genetic variation among human populations derives mainly from gradations in allele frequencies of subset of genes rather than from distinctive alleles present in specific populations. It is only through the accumulation of small allele-frequency differences across many loci that the genetic structure of a population, that is, the distinctive combination of allele frequencies, could emerge; see also below).

Several indices have been proposed to quantify the degree of genetic differentiation among two or more populations using series of gene frequencies. One such index is the F_{ST} distance, which measures the genetic variance between populations as a fraction of the total genetic variance. F_{ST} varies from 0 to 1. The closer F_{ST} is to 1, the higher is the genetic distance between two populations. This index has shown a high degree of correlation with other measures of genetic distances and since the data provided by Cavalli-Sforza are expressed in F_{ST} this index will be used in this study.

It is important to note that phenotypic characteristics (including anthropometric characters like stature or qualitative traits such as eye color or skin pigmentation) and the overall genetic structure of human populations are not related. For example, the pattern of overall genetic variation among populations differs substantially from traditional racial divisions (Figure 1). Morphologically similar peoples are not necessarily genetically similar overall. These findings confirm the reports by many investigators that physical anthropology data are not reliable to reconstruct past migrations. The most plausible reason is that external traits on which anthropometric studies are typically based on are particularly sensitive to natural selection. Only a very small fraction of the human genes are related to phenotypes that are under strong selection pressure (see for example Akey et al., 2002; Goldstein and Chikhi, 2002) . In contrast, most of the genes mentioned above, which differ between populations and are used to compute the genetic distance, are selectively neutral, that is, they lack selective advantage (see "Neutral Theory of Evolution", Kimura, 1968). As already clear to Darwin, neutral characters are best for reconstructing evolutionary history. If many genes used for the analysis show intercorrelated responses to the various environments in which human evolution has occurred, the measured genetic distance would be a reflection of the environments rather than of evolutionary history.

The absence of correlation between genetic distance and the color of the skin is particularly intriguing and would argue against a relationship between "cultural perception" and overall genetic features, as measured by Cavalli-Sforza et al., as well as by classical human population studies. Indeed, recent reports have suggested how skin pigmentation correlates with polymorphisms affecting single genes (Lamason et al., 2005; Soejima et al., 2006).

In contrast, a striking association between genetic distance and geography has recently emerged. Without using prior information about individual sampling locations, a clustering algorithm applied to multilocus genotypes from worldwide human populations produced genetic clusters largely coincident with major geographic regions (Rosenberg et al., 2005). For populations that are geographically close, genetic and geographic distances are often highly correlated, with genetic distance reaching an asymptote at about 1000-2600 miles on average (Figure 2).[7] Moreover, small discontinuous jumps in genetic distance are present for most population pairs on opposite sides of geographic barriers (Rosenberg et al., 2005), including in Europe, where sharp increases in genetic distance correspond to geographical impediments, including major mountains and seas (Barbujani and Sokal, 1990; see Figure 3).

III. GENETIC DISTANCE, GEOGRAPHY, AND TRADE

This section analyzes the relationship between genetic distance and geography. Our first goal is to show how geography has shaped genetic differences within Europe. Our starting point is Figure 3 (Sokal et al., 1990), which shows the main genetic changes within Europe. Sokal et al. (1990) identified 33 boundaries separating areas of sharp changes in gene frequencies across Europe and showed that the zones of abrupt genetic change in the European population correspond mostly to geographical boundaries. Specifically, the authors counted 22 physical boundaries, 4 mountainous and 18 marine. "In the 22 cases in which there are both physical barriers and genetic boundaries, it is reasonable to postulate that the causal arrow is likely to go more from physical barriers to both genetic and linguistic differentiation, rather than in other

[7] This issue is analyzed further in the empirical section, see in particular Table 1 for the geographic determinants of genetic distance.

directions" (Cavalli-Sforza, 1996, pag. 271).[8] Finally, as noted above, this notion is in line with classical genetic studies in humans and other organisms that also show a strong association between geographic boundaries and genetic distance.

In order to investigate more systematically the idea that genetic distance is conditioned by geographical factors, we run a regression with genetic distance as dependent variable and several geographical variables as control variables. The measure of genetic distance is derived from Cavalli-Sforza et al., p.270 (with F_{ST} derived from the analysis of the allele frequencies of 88 genes). The choice of geographical variables is suggested by the study of Sokal et al. (1990) and includes distance, number of mountains between countries, the presence of a common sea and average terrain elevation between two countries (as defined below), on the right-hand side. The results presented in the first three columns of Table 1 confirm that geographical measures and genetic distance among European countries are indeed correlated[9]. Table 1 and the literature reviewed above show that geography (including the distance between countries, the presence of major mountains chains, and common seas) plays a fundamental role in explaining genetic distance either by having determined past migration routes or by having separated populations, thereby contributing to the genetic drift.

Given the established correlation between geography and genetic distance, we hypothesize that geography affects both genetic distance and trade and that the correlation between trade and genetic distance is spurious. In the next section, we show that: i) the same geographic factors that affected genetic distance in the past are also important determinants of modern transportation costs; ii) in a standard gravity equation the impact of genetic distance on trade disappears once we introduce transport costs.

---

[8] Note also that some genetic boundaries cut countries in the middle (for instance in Germany, Italy, Finland, Island, Spain, and Greece.) This observation reinforces the point suggests that sharp genetic differences exist within national population, i.e. national borders are not always genetic borders. Moreover some abrupt changes follow linguistic and not geographical lines, for instance between Germany and the Netherlands. In other cases, there are linguistic barriers but no big genetic discontinuity (see for instance Germany and France, or Germany and Poland).

[9] Our measures of geography, particularly the presence of a common sea, are different from Sokal's measure. Sokal identified islands as geographic outliers, whereas in our case the presence of a common sea facilitates migration and hence reduces genetic distance. When we run a regression including island dummies among the regressors, the sign is positive and significant consistent with Sokal's findings.

A. Data

The bilateral export data are obtained from the United Nations COMTRADE database, the time span for our analysis is 1975-2002.[10] Our GDP data are obtained from the World Development Indicator of the World Bank; distance between capitals, common official language and contiguity dummies are obtained from a new dataset compiled at CEPII.[11]

We use two different sources of transport costs data[12]. The first measure (tc_{ij}) is taken from shipping company quotes collected by Import Export Wizard (IEW), a shipping company that provides estimates of transportation costs around the world. IEW calculates the surface freight data based on a survey of inter-modal and marine tariffs from carriers around the world. The variable tc_{ij} is the cost in U.S. dollars of transporting a "1000 kg unspecified freight type load (including machinery, chemicals, etc.) with no special handling required, using the optimal combination of going through land and water to transport the goods." The data refers to 2006. The advantage of this measure is that it represents the actual average transport costs and not some indirect measure or proxy, which are often plagued by measurement errors.

The second measure of transport costs is based on the "matched partner" technique (see Limao and Venables, 2000). Exporting countries report trade flows exclusive of freight and insurance (*fob*) and importing countries report flows inclusive of freight and insurance (*cif*). The ratio between *cif* and *fob* can then be used to construct an indirect measure of transportation costs, i.e. $itc_{ijt} = 1 - \frac{cif_{ijt}}{fob_{ijt}}$. The raw data come from the IMF's Direction of Trade Statistics from 1975 to 2002. The use of cif/fob data has been criticized because: 1) country statistics on cif and fob values are not always accurate; and 2) in many countries with trade restrictions and/or capital controls economic agents have perverse incentives and do not report the correct numbers; both caveats seem more relevant for less developed countries with poor administrative

---

[10] Prior to 2000, the data have been edited by Freenstra and Lipsey The data is available at http://cid.econ.ucdavis.edu/)

[11] The data is available at http://www.cepii.fr/anglaisgraph/bdd/distances.htm.

[12] For a review of the literature on transport costs see Anderson and van Wincoop (2004).

capacity and heavy controls, and arguably less relevant for Europe. Moreover, this measure is considered appropriate if it is used in combination with instrumental variables (see Hummels, forthcoming), which will be the focus of our analysis. Following Hummels (forthcoming), we restrict our analysis to *ad valorem* transportation costs which lie in a range between 0 and 100 percent, considered a reasonable range of variation. Finally, this measure allows us a panel analysis.

We construct a measure of geographical barriers using information on sea, mountain chains, and the average elevation of countries. We define a variable (mountains) identifying the number of major mountain chains between countries. According to the World Atlas, major mountain chains in Europe are: the Alps, the Apennines, the Atlantic Highlands (which include the Kjolen in Norway and Sweden, and the Pennines in the UK), the Balkan Mountains, the Massif Central, the Meseta, the Pyrenees, the Urals, the Carpathian Mountains and the Caucasus. We define a dummy "common sea" equal to one if a pair of countries share the same sea, which can be the Mediterranean, the Atlantic Ocean, or the Northern/Baltic Sea. Finally we construct a variable measuring the average elevation of countries which are between two trading partners. For instance, for the pair Germany-Italy this variable is equal to the average elevation of Germany, Austria, and Italy. This variable measures the difficulty of transportation/migration between countries. The sample statistics for the data are reported in Table A1.

B. Determinants of Transportation Costs and Genetic Distance

In this section we analyze whether the measures of geographic distance we have constructed have an impact on transportation costs and genetic distance. The importance of geography in determining transportation costs is well established; however, the standard measure of geography - log distance between (the capitals of) two countries - is only a first approximation of transport costs (see Hummels, 1998, Limao and Venables, 2000 and Eaton and Kortum, 2002).[13] A substantial literature in engineering has long observed that topographical characteristics such as terrain variability affect transportation costs (see Tsunokawa, 1983, and World Bank, 1987). For example, for the same horizontal distance, moving goods across variable

[13] See Clark, Dollar and Micco (2004) for a more nuanced measure of transport costs.

terrain requires more energy and time. Similarly, Limao and Venables (2000) and the World Bank (1998) find that non primary export high performers are island countries and none is landlocked. Our measure of geographic distance is a simple way to include geographical and topographical characteristics in the determination of transportation costs.

To understand the geographic determinants of transport cost, we run the following regressions:

$$\log(tc_{ij}) = \alpha_0 + \alpha_1 \log(D_{ij}) + \alpha_2 geography_{ij} + \alpha_3 country_i + \alpha_4 country_j + \varepsilon_{ijt}$$

where tc_{ij} are the transport cost, D_{ij} is the distance between *i* and *j*, *geography* is a vector with our proxies for geographical barriers between country *i* and country *j* (number of mountains, common sea, and average elevation of intermediate countries), $country_i$ and $country_j$ control for importer and trading countries fixed effects. The results of our regressions are reported in the last three columns of Table 1 for several combinations of geographical barriers. All variables have the expected sign. Distance, number of mountain chains, and average elevation between countries increase transportation costs, whereas the presence of a common sea reduces them. Our results are consistent with Limao and Venables (2000), who found that being landlocked significantly increase transportation costs, while being an island reduces them. Note that we use the same geographical determinants to explain genetic distance. As a robustness check, we also try the alternative measure of trade costs for 2002 (the measure based on "matched partner" technique, itc_{ijt}) and obtain analogous results (see Table A2 in the Appendix).

C. Genes versus Geography in Explaining Trade

In the previous section we found that the same geographical barriers that affect genetic distance are also crucial determinants of transportation costs. In this section, we show that when we instrument transportation costs with those measures of geographical characteristics that influenced also migration patterns in the past (and hence largely genetic distance nowadays), the impact of genetic distance on economic outcome (in this case trade) disappears. In order to do so, we estimate the following gravity equation:

$$\ln(X_{ijt}) = \beta_0 + \beta_1 gendist_{ij} + \beta_2 \ln(Y_{it}) + \beta_3 \ln(Y_{jt}) - \beta_4 \ln(D_{ij}) + \beta_5 C_{ij} + \beta_6 L_{ij} + \beta_7 E_{ijt} + \varepsilon_{ijt}$$

where X_{ijt} is the value of annual exports from country *i* to country *j*, $gendist_{ij}$ is the genetic distance between country *i* and *j* as defined by Cavalli-Sforza et al. (1994), Y_{it} is the real GDP of country *i*, D_{ij} is the distance between *i* and *j*, C_{ij} is a dummy variable for geographic contiguity between country *i* and *j*, L_{ij} is a dummy variable for common language between *i* and *j*, E_{ijt} is a dummy equal to one if country *i* and country *j* both use the Euro at time *t* and ε_{ijt} is the error term. We use a panel from 1975 to 2002, controlling for year and country of origin and country of destination fixed effects.

We begin by estimating a standard gravity equation. The first specification (Table 2, column 1) does not include genetic distance. Contiguity, distance and language have all the expected sign. The Euro variable is not significant. The importer's GDP has a negative and significant sign; since we include country of origin and country of destination fixed effects, the negative sign only captures the time variation of the series. As a next step, we introduce genetic distance among the regressors (specification 2). Genetic distance has negative sign and is significant at the 1 percent level. From this regression, one would be inclined to infer that cultural distance has a negative impact on trade. Specification (2) could be misleading, since genetic distance could capture the effect of omitted variables (transport costs). We argue that genetic distance is related to trade because it is a proxy for geographic impediments, which increase transportation costs between countries.

To address this issue we introduce transportation costs. The use of transport costs in an OLS context is not warranted because of clear problems of endogeneity (transport costs depend on the volume of trade, Hummels, 1999); therefore, we instrument transport costs with our measures of geography, which include the presence of major mountain chains, a common sea between countries, and the average elevation between countries. Note that we use the measure of indirect transport costs (itc_{ijt}) because it is the only available in a panel context. Columns 3-6 of Table 2 show the instrumental variable regressions. When we introduce transportation costs, genetic distance is not no longer significant, whereas the effect of transportation costs is negative and highly significant: a one-standard deviation increase in the log of transportation costs corresponds to an increase in trade of approximately 20% of the variation in the log of total exports. We report several alternative specifications. In column 3 we include only year effects, in column 4 we control for importer and exporter country fixed effects. Column 5 repeats the same specification of column 4, controlling in addition for country specific linear trends for both importing and exporting countries. As before, transportation costs appear to be crucial in the determination of trade, while genetic distance is once again not significant. In column 6, we instrument transportation costs using not only our measures of geography, but also log distance and contiguity (which are then not included in the gravity equation). Also with this alternative specification, transportation costs are still highly significant and the impact of genetic distance is nil.

As a robustness test, we repeat the same exercise using the alternative data on direct transportation costs from Import Export Wizard. The results are reported on Table 3 and are very similar to the panel specification. Genetic distance appears to be highly significant in the OLS regression, however its impact on trade is nil once transportation costs are introduced. A one standard-deviation increase in the log of transportation costs correspond to an increase in trade of approximately 14% of the variation in the log of total exports.

As a final step, we repeat the analysis for total exports but also using disaggregated data for differentiated products, reference priced and homogeneous goods (Table 4), following Rauch's classification (Rauch 1999).[14] When we replicate the analysis for these three different types of goods, the results still hold (genetic distance is not significant for any type of goods.) Moreover, the instrumented transportation costs are particularly relevant for homogeneous goods. Rauch (1999) constructs measures of transportability for different types of goods, finding a higher transportability for differentiated goods. Our results are consistent with his findings, as transportation costs matter more for homogeneous goods.

For all our specifications we run the appropriate tests (see Stock and Yogo (2002) and Moreira (2004)) to check that our instruments are weakly correlated with transportation costs.

[14] Homogeneous goods are goods traded or organized exchanges, reference priced goods are goods not traded on organized exchanges but nevertheless possessing a `reference price, differentiated goods are other commodities (Rauch, 1999). Over 60 percent of the value is in differentiated goods; about 20 percent belongs to reference priced goods, and the remaining to the organized trade.

The F statistics for the joint hypothesis that the instruments' coefficients are zero in the first stage regression, has always a significance level lower than 5% for all specifications.

IV. CULTURE VERSUS GENES: HINTS FROM AN INTERDISCIPLINARY APPROACH

In this section we explore the correlation between genetic distance and culture. Social scientists have speculated that genetic populations and culture are closely associated. This is the base for the current use of genetic distance as an instrument or a proxy for culture in economics. Being so crucial, the correlation between culture and genetic distance deserves a further analysis, after having explained the definition of cultural distance.

A. How to measure culture?

Culture is elusive to define and far more difficult to measure. Psychologists and anthropologists have attempted to provide scientific assessment of national characters. In social psychology, a national character is defined as a "relatively enduring personality character and patterns that are modal among the adult members of the society" (Inkeles and Levinson, 1969). In recent years, direct metrics of culture (independent of other variables) have gained wide acceptance in social psychology and anthropology. There are presently two widely used quantitative measures of personality and culture, allowing for quantitative research on cross-cultural studies: the five-factor model of personality (McCrae, 2002) and the IBM study dimensions of culture (Hofstede, 1980 and 2001).

The five factor model of personality, recently redefined as the "revised NEO personality inventory," is a model of trait structure in which relatively narrow and specific traits of personality are organized in terms of five broad factors: neuroticism, extraversion, openness to experience, agreeableness and conscientiousness.[15] Personality scales provide an obvious

[15] Neuroticism takes into account individual differences in the inclination to construct, perceive, and feel reality as being problematic. Extroversion refers to a tendency to seek contacts with the environment with enthusiasm and to live out positively. Openness groups together different types of behavior related to an active search of new experiences. Agreeableness refers more to the relational sphere of relationship with others (kindness, hostility, empathy). Conscientiousness focuses on issues such as orientation and control of impulses. Each of these five-factors is defined by 30 specific traits or facets (for example conscientiousness is represented by subscales measuring competence, order, dutifulness, achievement striving, self-discipline and deliberation). Among social scientists, it is widely accepted that the model is

(continued)

candidate to measure culture because mean levels of representative groups from different cultures can be compared quantitatively.

The IBM study of culture by Geert Hofstede (1980 and 2001) represents a widely accepted measure of cultural differences between countries. Hofstede divides cultures into five orthogonal dimensions: individualism (IDV), power distance (PD), uncertainty avoidance (UA), masculinity (MAS) and long term orientation (LTO).[16] Scores have been developed for several countries on these different dimensions. Since then, researchers have used the Hofstede measures to calibrate the different dimensions of a society's culture and then used the difference in these measures to capture the idea of "cultural distance." While Hofstede measures have not been free from criticism, they are by far the most likely used and cited cultural framework in international business, management, and applied psychology.

We calculate two measures of cultural distances, one based on the Five Factor Model of Personality and the other on the Hofstede measures, as follows:

$$\text{Cultural Distance}_{od} = \frac{\sqrt{\sum_{i=1}^{5}(C_{o,i} - C_{d,i})^2}}{5},$$

where $C_{o,i}$ is one of the dimensions of culture in one of the two models (i.e. IDV, PD, UA, MAS, LTO in Hofstede and N, E, O, A and C in the Five Factor model of Personality) for the country of origin and $C_{d,i}$ for the country of destination.

comprehensive-that is, it encompasses all major dimensions of personality and thus supersedes older trait models like Eysenck's (Eysenk & Eysenck, 1975) and Guilfrod's (Guilfrod & Zimmerman, 1976).

[16] Power distance focuses on the degree of equality, or inequality, between people in a given country. It represents the extent to which the less powerful members of organizations and institutions accept that power is distributed unequally. Individualism refers to the degree to which individuals are integrated into groups. Masculinity reflects the degree to which the society reinforces or not the traditional masculine-work role model of male achievement, control and power. The assertive pole has been called "masculine" and the caring pole "feminine". Uncertainty avoidance captures the society's attitude towards uncertainty, while long term orientation is associated with values such as thrift and perseverance as opposed to respect for tradition and fulfilling social obligation which are associated with short orientation.

B. Are measures of genetic and cultural distance correlated?

A bulk of literature has shown how the great majority of cultural differences among groups are consequences of cultural transmission rather than genetic differences; pronounced cultural differences exist among geographically contiguous peoples that are genetically indistinguishable and interbreed. Moreover, a recent study by Terraciano et al. (2005) has proved that, while individual differences in personality have a strong genetic foundation, cultural or national differences have no genetic basis. Even the hypothesis by Cavalli-Sforza et al (1994) that there is a strong correlation between linguistic groups and genetics distance has been questioned. For instance, Hunley and Long (2004) find that there is no relationship between language, as a part of the cultural heritage of a population, and the genetic structure in Native North Americans.

We also investigate the correlation between the available measure of culture and genetic distances within Europe. Table 5 shows bivariate correlations between the two measures of cultures and Cavalli-Sforza genetic distance measure. The cultural distances are positively and significantly correlated among themselves, but they are poorly or not significantly correlated with the measure of genetic distance. Our results are in agreement with previous findings.

V. POSSIBLE USES OF GENETIC DISTANCE IN ECONOMICS

If genetic distance cannot be used as a proxy for cultural distance, is there any other use in economics? In this section we discuss possible uses of genetic distance in economics. From the previous discussion it appears that the first obvious use of genetic distance is an instrument for transportation costs. Being exogenous and determined by geographical barriers, genetic distance could be simply used as a sophisticated measure for geographical impediments. In Table 6 we run a gravity equation using genetic distance as an instrument for transportation costs. As before, the first column is an OLS regression of a gravity equation with transportation costs among the regressors, controlling for fixed effects for both importing and exporting countries, as well as year dummies. As expected the coefficient on transportation costs is negative and significant. Column 2 uses genetic distance as an instrument for transportation costs. Column 3 includes country specific linear trends for both exporting and importing countries, finally column 4 has log distance and contiguity among the instruments. The first stage of this regression is reported in Table 7. As before, in the IV specification, not only the effect of transportation costs remains significant, but its magnitude increases substantially.

Besides being a proxy for geographical impediments, as a heritage of the past, the genetic make-up of a population can also provide valuable insights about the history of a country. Inasmuch history is believed to be an important determinant of current economic outcomes, genetics can add useful information. For instance, the genetic composition of a country, which was formerly a colony, can inform us on how much intermingling there was between colonizers and natives. Little intermingling will suggest that the colonizers' culture was simply imposed while more intermingling will indicate that population transfers also occurred.[17] Even more specifically, analysis of mitochondrial DNA and chromosome Y analysis can unveil the type of intermingling – if just male colonizers contributed or if whole nuclear families moved.

The genetic make-up of a population could also provide information on past environmental conditions, including the endemic diffusion of diseases such as malaria or the dietary habits, that shaped and acted upon specific phenotypes. Note that in this case, the measure of genetic distance proposed by Cavalli-Sforza, which is intentionally based on genes that likely do not affect phenotypes subject to natural selection, is not appropriate. Some of these applications are left for future research.

VI. CONCLUSIONS

This paper analyzes the validity of genetic distance as an instrument for cultural variables. We have shown that, despite the *prima facie* appeal of the variable genetic distance, there are several problems in the interpretation of the results. In particular, we have argued that: 1) genetic distance is not well correlated with cultural distance; 2) genetic distance is a better proxy for geographical variables.

---

[17] We credit Simon Johnson for this specific idea. Note that the idea that genetic analysis is a good instrument to understand the pattern of cultural spreading was the original motivation of Cavalli-Sforza et al. (1994). During the Neolithic revolutions in Europe new artifacts or techniques (or cultures) appeared to spread westwards. However, it is unclear if this spreading of culture was brought about by a spread of population or not. Cavalli—Sforza has argued that genetic analysis is compatible with the hypothesis of migration. Similarly, the transfer of techniques from colonizers to colonies may have happened with or without population intermingling and the subsequent economic outcome and institutions may depend on the way techniques have spread.

The paper provides an original contribution to the literature on transport costs. Several authors have shown that the simple measures of (log)-distance is only a first approximation for true transport costs (Hummels, Limao and Venables); in the context of the gravity models, many studies have included geographical variables such as insularity or contiguity to complement the standard crude measure of distance. Building on this tradition, we have shown that major mountains, common seas and countries elevation also contribute to transport costs. We finally show that being determined by major geographical impediments, genetic distance may be a good proxy for transportation costs.

References

Akey, J. M., Zhang, K., Jin, L., and Shriver, M.D. (2002) Interrogating a high-density SNP map for signatures of natural selection. Genome Res *12*, 1805-1814

Anderson, J. And E. van Wincoop (2004), "Trade Costs", Journal of Economic Literature, vol. 42, no. 3, 691-751

Cavalli-Sforza, L. L., and Feldman, M. W. (2003). The application of molecular genetic approaches to the study of human evolution. Nat Genet *33 Suppl*, 266-275.

Cavalli-Sforza, L., Menozzi, P. and A. Piazza (1994), *The History and geography of Human genes*, Princeton, Princeton University Press

Cavalli-Sforza, L. L., and Feldman, M. W. (2003). The application of molecular genetic approaches to the study of human evolution. Nat Genet *33 Suppl*, 266-275.

Eaton, J. and S. Kortum (2002), "Technology, Geography and Trade", Econometrica, Vol. 70 no. 5, 1741-1779

Clark, X., Dollar D. and A. Micco (2004), "Port efficiency, maritime transport costs, and bilateral trade", Journal of Development Economics, 417-450.

Goldstein, D. B., and Chikhi, L. (2002). Human migrations and population structure: what we know and why it matters. Annu Rev Genomics Hum Genet *3*, 129-152.

Guiso, L., Sapienza, P. and L. Zingales (2005), "Cultural Biases in Economic Exchange", NBER Working Paper #11005.

Hofstede, G. (1980), *Culture's Consequences*: *International Differences in Work-Related Values*. Beverly Hills: Sage Publication.

Hofstede, G. and R. McCrae (2004), "Personality and Culture Revised: Linking Traits and Dimension of Culture", Cross Cultural Research, 2004

Hummels, David, (1998), Towards a Geography of Transport Costs. Department of economics. University of Chicago. Mimeo.

Hummels, David, and Volodymyr Lugovskyy, Forthcoming, "Usable Data? Matched Partner Trade Statistics as a Measure of Transportation Costs." Review of International Economics.

Kimura, M. (1968). Evolutionary rate at the molecular level. Nature *217*, 624-626.

Lamason, R. L., Mohideen, M. A., Mest, J. R., Wong, A. C., Norton, H. L., Aros, M. C., Jurynec, M. J., Mao, X., Humphreville, V. R., Humbert, J. E., *et al.* (2005). SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science *310*, 1782-1786.

Lewontin, R.C., Rose L., and Kamin L. (1984). Not in Our Genes: Biology, Ideology and Human Nature. Pantheon, New York.

McCrae, R. and J. Allik (Eds.), 2002, *The Five-Factor Model of Personality across Cultures*, New York: Kluver Academic/Plenum Publisher

Putnam, R. (1993), *Making Democracy Work*, Princeton, Princeton University Press

Rosenberg, N. A., Mahajan, S., Ramachandran, S., Zhao, C., Pritchard, J. K., and Feldman, M. W. (2005). Clines, Clusters, and the Effect of Study Design on the Inference of Human Population Structure. PLoS Genet *1*, e70.

Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A., and Feldman, M. W. (2002). Genetic structure of human populations. Science *298*, 2381-2385.

Soejima, M., Tachida, H., Ishida, T., Sano, A., and Koda, Y. (2006). Evidence for recent positive selection at the human AIM1 locus in a European population. Mol Biol Evol *23*, 179-188.

Spolaore, E. and R. Wacziarg (2005), "The Diffusion of Development", mimeo

Terraciano. A et al. (2005), "National Character Does Not reflect Mean Personality Trait Levels in 49 Countries", *Science,* vol. 310, 7 October 2005, pp. 96-100.

Weber, M. 1958 *The Protestant Ethic and the Spirit of Capitalism*. Translated by T. Parsons. Scribner's Library,, New York.

Weber, Max 1958. *From Max Weber: Essays in Sociology*. Translated by H.H. Gerth and C. Wright Mills. Galaxy Book, New York, 1958

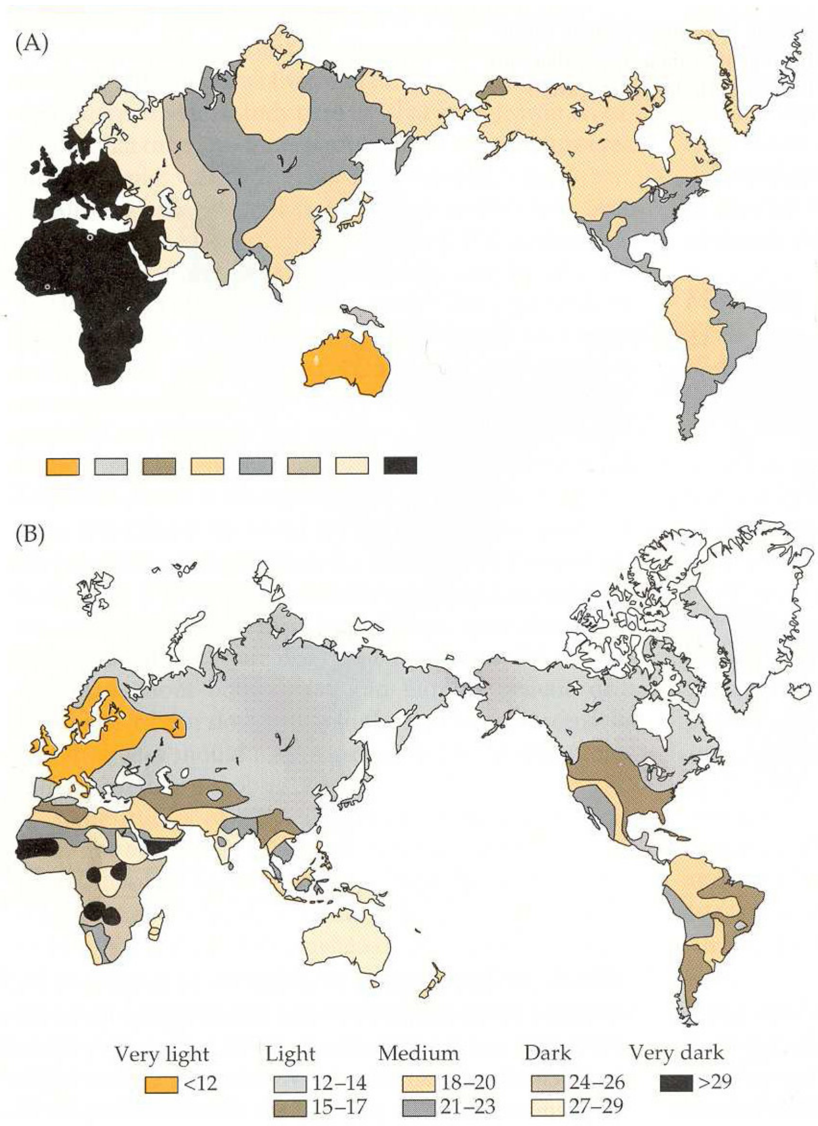


Figure 1. A. Division of the world's human population into eight classes of genetic similarity, based on overall difference and similarity at numerous enzyme and blood-group loci. The eight classes represented are arrayed in order of increasing difference.
B. Geographic distribution of skin color, classified in eight grades of pigmentation intensity (adapted from Cavalli Sforza et al., 1994)

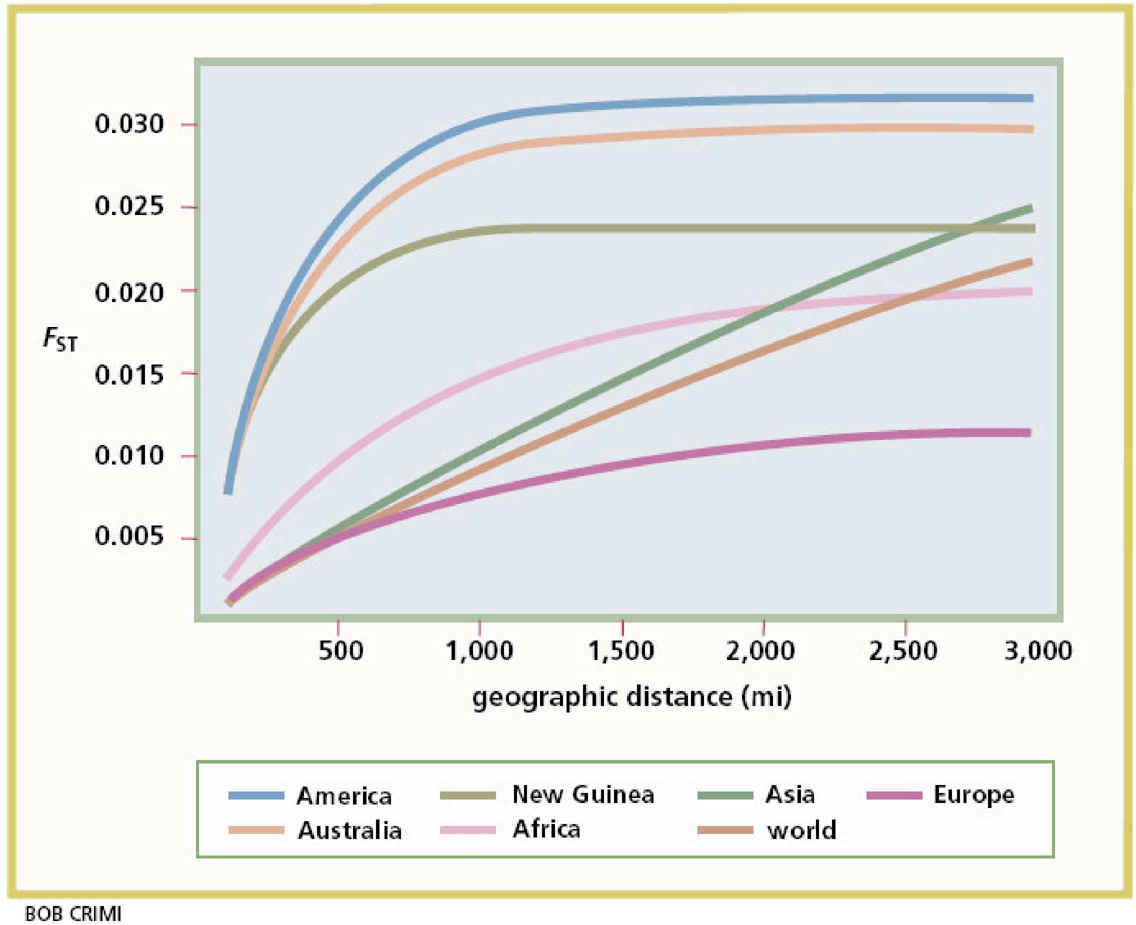


Figure 2

Relationship between Genetic and Geographic Distance (Cavalli-Sforza and Feldman, 2003)

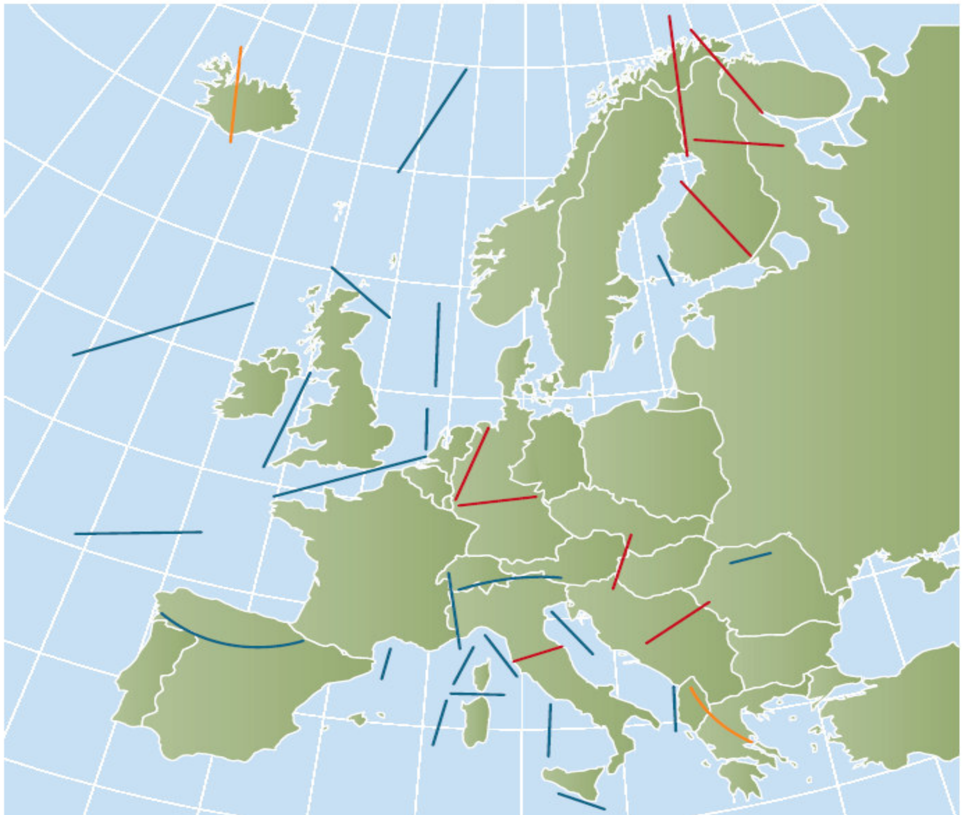


Figure 3

Zones of sharp genetic changes in Europe (Sokal, 1990)

Table 1.
Geographic Determinants of Genetic Distance and Transportation Costs

	Genetic Distance	Genetic Distance	Genetic Distance	$\ln(tc_{ij})$ Transport cost	$\ln(tc_{ij})$ Transport cost	$\ln(tc_{ij})$ Transport cost
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Log (distance)	21.92***	17.30***	17.30***	.0871***	.0880***	.0856***
	(3.104)	(3.725)	(3.729)	(.0036)	(.0067)	(.0039)
Number of mountain chains	1.007		-.8597	.0079***		.0069***
	(2.454)		(2.566)	(.0013)		(.0014)
Average elevation between countries		.0414**	.0442**		.0000***	.0000
		(.0214)	(.0225)		(.0000)	(.0000)
Common sea	-28.04***	-23.88***	-23.79***	-.0338***	-.0336***	-.0322***
	(6.723)	(7.512)	(7.562)	(.0013)	(.0042)	(.0041)
Country of origin fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country of dest. fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	470	470	470	1332	1332	1332
R-squared	0.95	0.95	0.95	0.93	0.93	0.93

* significant at 10%; ** significant at 5%, *** significant at 1%. Robust standard errors in parentheses

Table 2. Genetic Distance, Transportation Costs and Trade
Dependent Variable: Log Total Exports

	OLS (1)	*OLS* (2)	*IV* (3)	*IV* (4)	*IV* (5)	*IV* (6)
Transportation costs			-2.489***	-3.330***	-2.940***	-3.903***
(indirect transport costs)			(.2263)	(1.180)	(.9485)	(.4964)
Genetic Distance		-.0033***	-.0001	.0021	.0015	.0025
		(.0004)	(.0002)	(.0021)	(.0016)	(.0019)
Common Language	.5560***	.6613***	.2267	-.0123	-.1406	-.2253
	(.0567)	(.0480)	(.2250)	(.3629)	(.3425)	(.3591)
Euro	-.(.0415)	.0535	-.2543	-.3338	-.2365	-.3985
	0290	(.0396)	(.1762)	(.2878)	(.2656)	(.2951)
Log (distance)	-1.305***	-.8642***	.0022	-.2434	-.2487	
	(.0202)	(.0308)	(.0945)	(.2799)	(.2545)	
Contiguity	.2732***	.1931***	.2459**	-.1284	-.0549	
	(.0368)	(.0359)	(.1269)	(.2199)	(.1837)	
GDP exporter	.1718***	.7443***	.6496***	.6530	.5175	.5714
	(.0752)	(.1085)	(.0273)	(.4775)	(.7731)	(.5372)
GDP importer	-.3258***	-.8484***	.6713***	-1.734***	.2368	-1.843***
	(.0857)	(.1954)	(.0267)	(.5453)	(.8385)	(.5821)
Fixed year effects	yes	yes	yes	Yes	yes	Yes
Country dummies	yes	yes	no	Yes	yes	Yes
Country spec. linear trend	no	no	no	No	yes	No
Observations	20638	10096	7070	7070	7070	7070
R-Squared	0.85	.85	.96	.95	.96	.93

significant at 10, %, ** significant at 5%, *** significant at 1%. Robust standard errors in parentheses. Instruments include numbers of mountains and common sea. Column 6 also includes log(distance) and contiguity as instruments.

Table 3. Genetic Distance, Transportation Costs and Trade
Dependent Variable: Log Total Exports
Alternative measure of transportation costs

	OLS (1)	*OLS* (2)	*IV* (3)
Transportation costs (direct transport costs)			-15.204**
			(6.807)
Genetic Distance		-.0051***	.0010
		(.0013)	(.0029)
Common Language	.3525	.4172***	.7319***
	(.2460)	(.2062)	(.2701)
Euro	.0606	.1265	.1226
	(.1017)	(.1112)	(.1158)
Log (distance)	-1.467***	-.8089***	.6227
	(.0714)	(.0939)	(.6477)
Contiguity	.3232***	.3332***	.3693***
	(.1112)	(.1177)	(.1321)
Country dummies	Yes	Yes	Yes
Observations	1323	470	470
R-Squared	.90	.94	.92

significant at 10, %, ** significant at 5%, *** significant at 1%. Robust standard errors in parentheses.

Table 4. Genetic Distance, Transportation Costs and Trade
Dependent Variable: Log Total Exports

	Differentiated Goods			Reference Priced			Homogeneous Goods		
	OLS	*IV*	*IV*	*OLS*	*IV*	*IV*	*OLS*	*IV*	*IV*
Transportation costs		-3.891***	-3.497		-2.437***	-2.309***		-6.331***	-5.742***
(indirect trans. costs)		(1.365)	(.1143)		(.8527)	(.7426)		(2.360)	(1.890)
Genetic Distance	-.0025***	.0030	.0025	-.0038***	.0007	.0001	-.0028***	.0080**	.0063**
	(.0004)	(.0024)	(.0019)	(.0004)	(.0015)	(.0013)	(.0005)	(.0041)	(.0031)
Common Language	.6748***	-.1208	-.2859	.5313***	.0232	-.1205	.8153***	-.3204	-.6541
	(.0482)	(.4223)	(.4040)	(.0424)	(.2650)	(.2672)	(.0680)	(.6970)	(.6629)
Euro	.0159	-.4392	-.2927	.3638***	.0864	-.0147	.2688***	-.6337	-.3169
	(.0443)	(.3338)	(.3136)	(.0435)	(.2126)	(.2142)	(.0657)	(.5775)	(.5190)
Log (distance)	-.7476***	.0266	.0313	-.9975***	-.5915***	-.5464***	-1.328***	-.3077	-.2045
	(.0307)	(.3238)	(.2981)	(.0270)	(.2072)	(.2050)	(.0383)	(.5339)	(.4937)
Contiguity	.3030***	-.0718	.0108	.1490***	-.0440	.0084	.4005***	-.2498	-.1384
	(.0369)	(.2540)	(.2156)	(.0337)	(.1576)	(.1447)	(.0480)	(.4395)	(.3687)
GDP exporter	.9249***	-.1652***	.1909	.7571***	.7217**	-.3257	.0120	-.5939	-.4125
	(.1141)	(.6158)	(.9179)	(.1254)	(.3679)	(.6377)	(.1771)	(1.036)	(1.536)
GDP importer	-.7075***	.6964	.7152	.3917***	.0787	.4801	.5474***	-.9245	-.4765
	(.1921)	(.5554)	(.9620)	(.1258)	(.3818)	(.6423)	(.1922)	(.9622)	(1.607)
Fixed year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country specific time effects	No	no	Yes	No	No	Yes	No	No	Yes
Observations	10070	7054	7054	9943	6960	6960	9579	6735	6735
R-Squared	.87	.93	.94	.88	.97	.97	.77	.72	.78

* significant at 10%; ** significant at 5%, *** significant at 1%. Robust standard errors in parentheses.

Table 5.
Bivariate Correlations between Cultural Distances and Genetic Distance

	Cavalli-Sforza Genetic Distance	Hofstede Cultural Distance	NEO- P I Distance
Cavalli-Sforza Genetic Distance	1		
Hofstede Cultural Distance	0.15**	1	
NEO- P I Distance	-0.00	0.57***	1

Correlation matrix of the cultural and genetic distance variables.
*, **, *** denote significance at 10, 5 and 1 percent level, respectively.

Table 6
Genetic Distance, Transportation Costs and Trade
Dependent Variable: Log Total Exports
Instrumenting Transportation Costs with Genetic Distance

	OLS *(1)*	*IV* *(2)*	*IV* *(3)*
Transportation costs	-1.249***	-1.811***	-1.754***
(indirect transport cost)	(.0098)	(.5742)	(.5890)
Common Language	.3858***	.2595	.1522
	(.0706)	(.1995)	(.2151)
Euro	-.0597	-.1432	-.1098
	(.0504)	(.1591)	(.1671)
Log (distance)	-1.308***	-.5419***	-.5193***
	(.0252)	(.1691)	(.1846)
Contiguity	.1957***	.0227	.0562
	(.0476)	(.1195)	(.1140)
GDP exporter	.2797***	.8034***	.6173
	(.0878)	(.2740)	(.4929)
GDP importer	-.5246***	-1.510***	.2766
	(.100)	(.3597)	(.5793)
Fixed year effects	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes
Country spec. linear trend	No	No	Yes
Observations	14263	7070	7070
R-Squared	.83	.98	.98

* significant at 10%; ** significant at 5%, *** significant at 1%.
Robust standard errors in parentheses.

Table 7
IV Regression, First Stage
Genetic Distance as an Instrument for Transportation Costs

Genetic Distance	.0014***
	(.0003)
Common Language	-.1790***
	(.0884)
Euro	-.1254**
	(.0649)
Log(Distance)	.1965***
	(.0370)
Contiguity	-.0995**
	(.0496)
GDP exporter	-.0989
	(.1323)
GDP importer	-.1478
	(.1344)
Observations	7070
R-squared	.20

* significant at 10%; ** significant at 5%, *** significant at 1%.
Robust standard errors in parentheses.

Table A1
Summary Statistics

	Mean	Std. Dev.	Observations
Log Distance	7.134	.6365	1480
Contiguity	.0959	.2946	1480
Common Language	.0135	.1154	1480
Genetic Distance	99.22	115.24	470
Islands	.1277	.3338	1480
Landlocked	.1540	.3611	1480
Euro	.0235	.1517	26460
Number of mountains	.9202	.9420	1480
Number of shared rivers	.0770	.2955	1480
Average elevation b/w countries	159.08	122.04	1474
Common sea	.3594	.4800	1480
Log(GDP) Exporter	25.22	1.85	22280
Log(GDP) Importer	25.22	1.85	22280
Log(Total Exports)	11.29	2.74	25872
Log Transp. Costs (1-cif/fob) (indirect transport costs)	2.70	.9423	17090
Log Transp. Costs (direct transport costs)	5.27	.1603	1332

Table A2
Geographic Determinants of Transportation Costs
[(transportation costs=(1-cif/fob)]

	ln (itc_{ij2004}) (1)	Ln (itc_{ij2004}) (2)	ln (itc_{ij2004}) (3)
Log (distance)	.1289***	.1237***	.1213***
	(.0433)	(.0448)	(.0464)
Number of mountain chains	.0112		.0065
	(.0277)		(.0286)
Average elevation between countries		.0001	.001
		(.0002)	(.0002)
Common sea	-.1271**	-.1184*	-.1173*
	(.0564)	(.0633)	(.0631)
Country of origin fixed effects	Yes	Yes	Yes
Country of dest. fixed effects	Yes	Yes	Yes
Observations	1120	1120	1120
R-squared	.22	.22	.22

* significant at 10%; ** significant at 5%, *** significant at 1%.
Robust standard errors in parentheses.

Table A3- Genetic Distance, Transportation Costs and Trade
Alternative specifications
Dependent Variable: Log(Total Export)

	IV (1a)	*IV (2a)*	*IV (3a)*	*IV (4a)*	*IV (1b)*	*IV (2b)*	*IV (3b)*	*IV (4b)*
Transportation costs	-2.840***	-2.146***	-2.109***	-3.580***	-2.525***	-2.278***	-2.098***	-3.405***
	(.2659)	(.5191)	(.5234)	(.4436)	(.2164)	(.4297)	(.3944)	(.3750)
Genetic Distance	-.0000	.0004	.0004	.0015	-.0001	.0006	.0004	.0010
	(.0002)	(.0011)	(.0011)	(.0017)	(.0002)	(.0011)	(.0010)	(.0016)
Common Language	.1723	.1997	.0647	-.1200	.2212	.1759	.0673	-.0631
	(.2540)	(.2207)	(.2258)	(.3263)	(.2279)	(.2226)	(.2059)	(.3032)
Euro	-.3128	-.1852	-.1477	-.3565	-.2602	-.2018	-.1465	-.3339
	(.2006)	(.1755)	(.1874)	(.2706)	(.1786)	(.1792)	(.1823)	(.2558)
Log (distance)	.0980	-.4761***	-.4385***		.0119	-.4500***	-.4409***	
	(.1111)	(.1416)	(.1529)		(.0949)	(.1357)	(.1320)	
Contiguity	.2355*	-.0105	.0230		.2449**	-.0237	.0240	
	(.1432)	(.1279)	(.1238)		(.1283)	(.1340)	(.1217)	
GDP exporter	.6420***	.7703***	.5872	.5986	.6466***	.7571***	.5882	.6133
	(.0295)	(.3137)	(.5719)	(.4914)	(.0269)	(.3294)	(.5697)	(.4712)
GDP importer	.6200***	-1.559	.2647	-1.802	.6685***	-1.579***	.2654	-1.779***
	(.0312)	(.3863)	(.6517)	(.5413)	(.0258)	(.4000)	(.6497)	(.5203)
Fixed year effects	Yes	yes	yes	Yes	yes	yes	Yes	yes
Country dummies	No	yes	yes	Yes	No	yes	Yes	yes
Country spec. linear trend	No	no	yes	No	No	no	Yes	no
Observations	7070	7070	7070	7070	7070	7070	7070	7070
R-Squared	.96	.98	.98	.94	.97	.97	.98	.95

Significant at 10%; ** significant at 5%, *** significant at 1%. Robust standard errors in parentheses
Columns a includes as instruments average elevation between countries and common sea, while columns b include number of mountain chains, average elevation between countries and common sea.

Table A4
IV Regression, First Stage
Geographical Barriers as a Measure for Transportation Costs

	Log (Transp. Costs)	Log (Transp. Costs)	Log (Transp. Costs)
Number of mountain chains		.0897***	.0691***
		(.0146)	(.0154)
Average Elev. b/w countries	.0007***		.0005***
	(.0001)		(.0001)
Common sea	-.1176***	-.1371***	-.0901***
	(.0260)	(.0241)	(.0266)
Common language	-.1971**	-.1921**	-.1740**
	(.0875)	(.0875)	(.0875)
Euro	-.1414**	-.1634***	-.1595***
	(.0622)	(.0623)	(.0622)
Genetic Distance	.0002***	.0002**	.0001**
	(.0000)	(.0000)	(.0000)
Log(distance)	.2268***	.2309***	.1958***
	(.0271)	(.0266)	(.0279)
Contiguity	.0530	-.0211	.0143
	(.0454)	(.0454)	(.0462)
GDP exporter	-.0895***	-.0886***	-.0940***
	(.0079)	(.0079)	(.0080)
GDP importer	-.0894***	-.0887***	-.0940***
	(.0079)	(.0079)	(.0080)
Observations	7070	7070	7070
R-squared	.11	.11	.12

* significant at 10%; ** significant at 5%, *** significant at 1%. Robust standard errors in parentheses.