Is there a role for genetics in economic development?

Luis Angeles*

January 24, 2012

Abstract

Spolaore and Wacziarg (2009) have presented evidence supporting a role of genetic distance to the United States as a barrier to economic development. We extend their empirical work by controlling for the share of Europeans and European descendants in the population. We find that the role of genetic distance disappears and offer two alternative interpretations of the patterns in the data.

Keywords: Genetics; economic development; European settlement.

JEL classification: O1, O3, O4.

The economics literature has begun to take advantage of genetic data in order to explain different economic phenomena.[1] A particularly intriguing possibility is that genetic differences may play a role in explaining differences in economic development, an idea that has been advanced by Spolaore and Wacziarg (2009) - henceforth SW. These authors show that a country's genetic distance to the world technological leader, the United States, is a powerful predictor of its income per capita. This note qualifies the results of SW by subjecting them to some additional empirical testing.

---

*Economics, University of Glasgow. Adam Smith Building, Glasgow G12 8RT, United Kingdom. Email: luis.angeles@glasgow.ac.uk . Tel: +44 141 330 8517.

[1] See, for instance, Guiso et al. (2009), Giuliano et al. (2006) and Desmet et al. (2009).

SW have been careful to point out that "our results provide *no* evidence for a direct effect of specific genes on income or productivity" (p. 470, italics in the original). Their mechanism relies instead in an assumed relationship between genetic and cultural distance. As they point out, genetic distance is a good proxy for the time since two populations have diverged from a single, parent population. The authors then assume that the longer the time since divergence the greater the chances that these populations have developed different cultural practices (or, more generally, different intergenerationally transmitted traits, which may be cultural or biological). The final piece of the mechanism is that cultural differences act as "barriers to the diffusion of development" because "more closely related societies are more likely to learn from each other and adopt each other's innovations" (Spolaore and Wacziarg 2009, p. 470). Thus, genetic/cultural distance with respect to the world technological leader would make economic development difficult.

SW do not offer any direct evidence of their hypothesized links between genetic and cultural distance or between cultural distance and technological adoption or economic success. Instead, the authors focus on the reduced-form relationship between genetic distance to the United States and income per capita and present a large array of econometric results that support their story. Most of their results are obtained using country pairs as the unit of analysis and controlling for alternative explanations of long-run economic development such as geography, climate, colonial past, and measurable aspects of culture such as language and religion.[2]

SW are worthy of much praise for their innovative use of genetic data and their careful handling of the empirical analysis. The economics literature will probably gain much from further extending in the direction marked by these and other authors. We will, however, beg to differ with the overall conclusions of SW. As we argue below, a simple extension to their empirical framework shows their results are less robust than at first sight.

---

[2] For some attempts at measuring cultural differences beyond language and religion see Fearon (2003), Linders et al. (2005) and Tadesse and White (2010a, 2010b).

***

The genetic distance between two given countries in the world can be predicted to a large extent by answering two questions. First, is either of them an African country? Second, to what extend is their population of European descent?

The first question sets African and non-African countries aside. This makes sense because non-Africans are a subgroup of Africans in terms of genetic diversity, a consequence of the pattern of evolution and diffusion of homo sapiens throughout the world.[3] As a consequence, the genetic distance between two non-African countries is typically smaller that the genetic distance between an African and a non-African country.

The second question recognizes the ubiquity and importance of the European colonial expansion over the last few centuries. Colonialism brought large number of European settlers abroad, modifying and in some cases completely changing the genetic composition of some nations - the United States among them.[4]

It is thus the case that being an African country is a powerful predictor of a large genetic distance vis-a-vis the Untied States while having a large European population is a powerful predictor of a small genetic distance to this same country. The importance of these two factors can be appreciated in table 1, where they are used as explanatory factors of the genetic distance

---

[3] The spread of modern humans out of Africa took place in migratory waves which gave rise to the "serial founder effect" of population genetics. In short, only a small sample of the initial African population migrated to Eurasia where new virgin lands allowed it to multiply its numbers and create a whole new population. This initial migratory group contained only a subsample of the genetic variation within Africa, resulting in a much reduced level of genetic diversity in Eurasia as compared to Africa. For a similar reason, Amerindians can be seen as a subgroup of Eurasians in terms of genetic diversity.

[4] See Putterman and Weil (2010) for a full account of population movements between countries since 1500. As is clear from their figure 2, Europeans dominated such movements. For other consequences of European settlement on present-day socioeconomic outcomes see Angeles (2007) on income inequality and Angeles and Neanides (2009, 2011) on aid effectiveness and corruption.

Table 1
Explaining genetic distance

Dependent variable: Genetic distance to the United States	(1)	(2)	(3)
Sub-Saharan Africa dummy	0.082 (0.007)**		0.064 (0.007)**
Share of Europeans and European descendants in the population		-0.076 (0.006)**	-0.050 (0.005)**
Observations	180	178	178
R^2	0.48	0.39	0.63

Note: Standard errors are given in parenthesis. The symbols * and ** denote statistical significance at the 5% and 1% level.

of each country to the United States in simple OLS regressions. Not only are these two factors statistically significant predictors at the 1% level, but together they explain 63% of the variation in genetic distance across the world.[5]

***

To their credit, SW subject their results to the inclusion of a dummy variable for sub-Saharan African countries. This is a reasonable test to make because many things set African countries aside other than a large genetic distance to the United States. SW mainly refer to Diamond (1997), who masterly documented the biological and geographic handicaps that hindered Africa's transition from hunther-gathering to agriculture. Additional factors contributing to Africa's long-standing economic backwardness are the high prevalence of infectious diseases (McNeill 1976, Gallup et al. 1999) and the many centuries of the slave trade (Nunn 2008).

SW find that their qualitative results continue to hold when controlling for African countries, although the magnitude of the estimated effect of

---

[5] Please refer to the appendix for data sources and definitions, including an explanation of the measure of genetic distance.

genetic distance falls by about one third (tables IV and V in their paper). They interpret this result as evidence that "Diamond's hypothesis on the long term diffusion of economic development is complementary to ours" (p. 504).

What SW do not control for, however, is the prevalence of Europeans or European descendants within the population. One may understand why, had they been aware of the importance of this variable in explaining genetic distance, they would have chosen not to include it in their regressions. A large share of Europeans in the population and a close genetic distance to the United States are closely related concepts. It may well be that the share of Europeans in the population is positively related to economic development precisely because of Spolaore and Wacziarg's story: small genetic distance, small cultural distance, smooth transfer of ideas.

To this argument we have two answers. First, it may also be that the success of countries of European population is explained by factors other than culturally-facilitated technological transmission. European settlers had higher levels of human and physical capital and had already acquired a familiarity with modern European technology by the time they went abroad. They also took their laws and institutions with them, which could result in economic development independently of technological transmissions.[6] If these were the mechanism at work then it would make sense to control for Europeans and European descendants when estimating the effects of genetic distance.

The second answer is even simpler. Let us accept, in accordance with Spolaore and Wacziarg, that countries with an European population are rich because their close genetic and cultural distance with the United States (and Europe) allowed them to use their ideas. We would still want to know whether this mechanism is in place for countries with no European population. According to table 1, we would still have 37% of the variation in genetic distance to the United States once we control for sub-Saharan Africa and for

---

[6] See Acemoglu et al. (2001), Glaeser et al. (2004) and the related literature.

the share of Europeans and their descendants in the population. The genetic characteristics of Asia, North Africa and the Pacific are not explained by either of these two variables. Is the thesis of Spolaore and Wacziarg relevant for them?

***

As we mentioned above, most of the results in SW are obtained using country pairs as the unit of analysis. This approach increases the number of available observations from about 150 in a standard cross-country setting to 10,000 or more observations of country pairs. We will thus focus on their preferred empirical specification, which we reproduce below:

$$|\log y_i - \log y_j| = \gamma_0 + \gamma_1 G^R_{ij} + \gamma_2 X_{ij} + \nu_{ij} \tag{1}$$

In equation (1) $|\log y_i - \log y_j|$ is the difference, in absolute value, of the logarithm of income per capita between countries i and j, G^R_{ij} is the relative genetic distance with respect to the United States and X_{ij} a set of control variables measuring alternative determinants of long-run development. If we denote the genetic distance between countries i and j by G^D_{ij}, the relative genetic distance between two countries with respect to the United States is defined as $G^R_{ij} = \left| G^D_{i,US} - G^D_{j,US} \right|$.

Our set of control variables will take into account the effects of geography, climate and some aspects of culture - all of which have been considered by SW in one way or another. We include the absolute latitude of a country; a dummy for landlocked countries; dummies for former colonies of Britain, France, Spain or Portugal, and any other country; the share of English, French, and Spanish or Portuguese speakers; and the share of the population professing the Protestant, Catholic and Muslim religions. All these variables are included after taking the absolute value of the difference between countries i and j. Note that for a dummy variable like being landlocked the result is a new dummy taking a value of 0 if either both countries are landlocked or none of them is landlocked, and a value of 1 otherwise. Such variable would capture the income distance resulting from a difference in that particular trait.

Our results are summarized in table 2. The first two columns of this table reproduce the central finding of SW: relative genetic distance to the United States has a positive, sizeable and statistically significant effect on income differences when used as the sole regressor (column 1) and with an important set of control variables (column 2). The estimated coefficient of relative genetic distance that we obtain is within the range of results reported by these authors. A coefficient of 5.123, as in column 2, implies that a one standard-deviation change in relative genetic distance is associated with a change in the income gap of 25%.

Column 3 adds a dummy for sub-Saharan Africa (after taking differences and absolute value as indicated above). The coefficient for this variable takes the expected positive sign (pairs with one country from sub-Saharan African tend to have a larger income difference) and is statistically significant. The coefficient on relative genetic distance falls by about half with respect to its previous value but remains large and statistically significant. This result is consistent with those obtained by SW in similar regressions and shows that their thesis survives the hypothesis of African low incomes being caused by factors unrelated to genetic distance.

Column 4 then adds the share of Europeans and European descendants in the population. We also include a dummy variable for former USSR and Warsaw pact countries in order to distinguish between the well-known income-reducing effect of communism from the otherwise income-increasing effect of having a European population. This is important since a large number of European countries lived through much of the 20th century under a communist regime.

The results of column 4 are noteworthy. As expected, we find a positive coefficient on the share of Europeans in the population - large differences

Table 2
Country pairs regressions

Dependent variable: log of GDP per capita in 1995

	(1)	(2)	(3)	(4)
Relative genetic distance to the United States	6.793 (0.252)**	5.123 (0.261)**	2.297 (0.296)**	0.510 (0.299)
Sub-Saharan Africa			0.546 (0.027)**	0.578 (0.026)**
Share of Europeans and European descendants				0.687 (0.033)**
Former USSR or Warsaw Pact				-0.599 (0.031)**
Latitude		0.012 (0.001)**	0.013 (0.001)**	0.007 (0.001)**
Landlocked		0.306 (0.024)**	0.293 (0.023)**	0.390 (0.023)**
Former British colony		-0.046 (0.023)*	-0.058 (0.022)**	-0.068 (0.022)**
Former French colony		0.059 (0.026)*	-0.065 (0.027)*	-0.101 (0.026)**
Former Spanish or Port. colony		-0.121 (0.034)**	-0.171 (0.033)**	-0.152 (0.032)**
Other former colonies		-0.02 (0.042)	-0.069 (0.042)	-0.034 (0.040)
Share of English speakers		0.239 (0.034)**	0.281 (0.032)**	0.249 (0.033)**
Share of French speakers		0.331 (0.064)**	0.367 (0.068)**	0.213 (0.067)**
Share or Spanish/Port. speakers		-0.098 (0.039)*	-0.036 (0.038)	-0.164 (0.038)**
Share of Protestants		0.438 (0.053)**	0.504 (0.052)**	0.423 (0.053)**
Share of Catholics		0.168 (0.039)**	0.19 (0.039)**	0.161 (0.039)**
Share of Muslims		0.197 (0.031)**	0.186 (0.031)**	0.174 (0.031)**
Observations	13861	13203	13203	12880
R^2	0.06	0.11	0.14	0.18

Note: Standard errors are given in parenthesis. The symbols * and ** denote statistical significance at the 5% and 1% level.

in this variable go hand in hand with large income differences. More remarkable, however, is that in this regression the effect of genetic distance on economic development is an order of magnitude smaller than in the initial regressions of columns 1-2 and is no longer statistically significant.

A moment of reflection shows that this makes perfect sense. The genetic variation that is left is largely given by the countries of Asia, North Africa, and the Pacific. Genetic data shows that, in accordance with the pattern of diffusion of early humans around the globe, geographic distance from Europe is a very good predictor of genetic distance with respect to Europeans (and thus with respect to the United States). It follows that within this group Middle Easterners and Northern Africans are genetically closest to the United States, followed by people from South Asia, Central Asia, and finally East Asia and the Pacific. Or, contrary to what genetic distance would predict, it is well East Asian countries that have experienced the most economic development within this group, while North Africa and South Asia have lagged behind. In short, the thesis of SW does not seem to apply to this important group of countries.

We bring some further support to the case by repeating the above exercise using traditional cross-country regressions (similar to table 1 in SW). In this case, the GDP per capita of each country is predicted to be negatively related to its genetic distance with respect to the United States. As table 3 demonstrates, results parallel those obtained using country pairs. Once again genetic distance appears to have a strong effect on income levels, an effect that remains in place when we control for sub-Saharan African countries. Once the share of Europeans and European descendants is included, however, genetic distance to the United States appears to be of no importance.

The results of this note may thus be summarized as follows. From a statistical perspective, the relationship between genetic distance to the United States and economic development uncovered by Spolaore and Wacziarg (2009) disappears when we add the share of Europeans and their descen-

Table 3
Cross-country regressions

Dependent variable: log of GDP per capita in 1995

	(1)	(2)	(3)	(4)
Genetic distance to the United States	-16.276	-10.119	-6.326	-0.780
	(1.765)**	(2.770)**	(3.135)*	(3.337)
Sub-Saharan Africa			-0.784	-1.205
			(0.308)*	(0.351)**
Share of Europeans and European descendants				1.144
				(0.402)**
Former USSR or Warsaw Pact				-1.73
				(0.274)**
Latitude		0.027	0.027	0.025
		(0.010)**	(0.010)**	(0.012)*
Landlocked		-0.892	-0.859	-0.634
		(0.224)**	(0.217)**	(0.186)**
Former British colony		0.53	0.638	0.350
		(0.337)	(0.333)	(0.340)
Former French colony		-0.26	-0.007	-0.300
		(0.318)	(0.312)	(0.333)
Former Spanish or Port. colony		-0.434	-0.275	-0.375
		(0.313)	(0.300)	(0.318)
Other former colonies		-0.273	-0.17	-0.425
		(0.446)	(0.423)	(0.462)
Share of English speakers		0.407	0.157	0.144
		(0.316)	(0.359)	(0.341)
Share of French speakers		-0.097	-0.133	-0.395
		(0.663)	(0.894)	(0.836)
Share or Spanish/Port. speakers		0.145	-0.131	-0.400
		(0.326)	(0.347)	(0.336)
Share of Protestants		1.852	1.82	0.908
		(0.500)**	(0.512)**	(0.498)
Share of Catholics		1.466	1.544	1.197
		(0.397)**	(0.399)**	(0.360)**
Share of Muslims		0.092	0.126	0.365
		(0.396)	(0.390)	(0.404)
Observations	167	163	163	161
R^2	0.29	0.60	0.62	0.69

Note: Standard errors are given in parenthesis. The symbols * and ** denote statistical significance at the 5% and 1% level.

dants in the population as an additional control. This may be interpreted in two alternative ways. First, the mechanism emphasized by SW is present in the data and explains the good economic performance of countries with a European population. But the mechanism is of no relevance to non-African countries without a European population. Second, the mechanism emphasized by SW is not present in the data - the good economic performance of countries with a European population is due to factors unrelated to genetic distance. Distinguishing between these two interpretation would require the use of measures correlated with the degree of Europeans in the population yet unrelated with genetic distance to the United States. This challenging task is beyond the scope of the present note and is left for future research.

References

Acemoglu, D., Johnson, S. and Robinson, J. A. 2001, The colonial origins of comparative development: an empirical investigation, American Economic Review 91 (5), p. 1369-1401.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S. and Wacziarg, R. 2003, Fractionalization, Journal of Economic Growth 8, 155-194.

Angeles, L. 2007, Income Inequality and Colonialism, European Economic Review 51 (5), 1155-1176.

Angeles, L. and Neanidis, K. 2009, Aid Effectiveness: The Role of the Local Elite, Journal of Development Economics 90 (1), 120-134.

Angeles, L. and Neanidis, K. 2011, Colonialism, Elite Formation and Corruption, Discussion Paper 2011-02, Economics, University of Glasgow.

Cavalli-Sforza, L. L., Menozzi, P. and Piazza, A. 1994, The History and Geography of Human Genes, Princeton: Princeton University Press.

Desmet, K., Le Breton, M., Ortuño-Ortin, I. and Weber, S. 2009, Stability of Nations and Genetic Diversity, mimeo University Carlos III.

Diamond, J. 1997, Guns, Germs and Steel, New York: W. W. Norton.

Fearon, J. D. 2003, Ethnic and cultural diversity by country, Journal of Economic Growth 8, 195-222.

Gallup, J. L., Sachs, J. and Mellinger, A. 1999, Geography and Economic Development, CID working paper n.1, Harvard University.

Giuliano, P., Spilimbergo, A. and Tonon, G. 2007, Genetic, Cultural and Geographic Distances, mimeo Harvard University.

Glaeser, E. L., La Porta, R., Lopez de Silanes, F. and Shleifer, A., 2004, Do Institutions cause growth?, Journal of Economic Growth 9, 271-303.

Guiso, L, Sapienza, P. and Zingales, L. 2009, Cultural biases in economic exchange?, Quarterly Journal of Economics 129 (August 2009), 1095-1131.

La Porta, R., López-de-Silanes, F., Shleifer, A. and Vishny, R. 1999, The Quality of Government, Journal of Law, Economics and Organization 15 (1), 222-279.

Linders, G. J., Slangen, A., de Groot, H. L. F. and Beugelsdijk, S. 2005, Cultural and Institutional determinants of bilateral trade, Tinbergen Discussion Paper 05-074/3.

McNeill, W. H. 1976, Plagues and peoples, Oxford: Basil Blackwell.

Nunn, N. 2008, The Long Term Effects of Africa's Slave Trades, Quarterly Journal of Economics 123 (1), 139-176.

Putterman, L. and Weil, D. N. 2010, Post-1500 population flowsand the long-run determinants of economic growth and inequality, Quarterly Journal of Economics 130, 1627-1682

Spolaore, E. and Wacziarg, R. 2009, The Diffusion of Development, Quarterly Journal of Economics 129 (May 2009), 469-529.

Tadesse, B. and White, R. 2010a, Cultural distance as a determinant of bilateral trade flows: Do immigrants counter the effects of cultural differences? Applied Economic Letters 17, 147-152.

Tadesse, B. and White, R. 2010b, Does cultural distance hinder trade in goods? A comparative study of nine OECD member nations, Open Economies Review 21, 237-261.

Appendix: definitions and data sources

GDP per capita

From the World Bank's World Development Indicators (Edition 2010). Real GDP per capita in US$ of the year 2000, for the year 1995 (in accordance with Spolaore and Wacziarg 2009).

Genetic distance

From Spolaore and Wacziarg (2009), who obtain the data from Cavalli-Sforza et al. (1994). These authors use the measure of genetic distance know as F_{ST}, which can be described as the fraction of total heterozygosity in a group of two populations that cannot be explained by the average within-group heterozygosity.

The heterozygosity of a given gene in a population is the probability that, if we pick up two random observations of that gene in the population, they will present different forms of that gene (what is known as the different alleles of the gene). Because every person has two versions of every gene (one from each parent), this measure is also equal to the proportion of people with different alleles for their two genes (assuming, of course, that mating is random with respect to that gene). The measure is then averaged over a certain number of genes to obtain the overall heterozygosity in a population (Cavalli-Sforza et al. 1994).

Share of Europeans and European descendants in the population

From Alesina et al. (2003). We consider as Europeans all groups who belong to the ethnic majority of a European country (e.g. "Germans", "Bulgarians" and so on) together with broader classifications such as "western Europeans", "eastern Europeans" or, simply, "Europeans". We do not consider Turkey and Central Asia as parts of Europe, while all former republics of the Soviet Union west of the Ural mountains are counted in. In non-European countries, we count as European descendants the groups described as "whites". We do not count mestizos or people of other mixed ethnic backgrounds as European descendants.

Shares of English, French, and Spanish or Portuguese speakers
From Alesina et al. (2003).

Shares of Protestants, Catholics and Muslims (for the year 1980), Latitude, Landlocked
From La Porta et al. (1999).

Colonial dummies, dummy for former USSR and Warsaw pact
Constructed by the author.