LONG-TERM BARRIERS TO ECONOMIC DEVELOPMENT

Enrico Spolaore
Romain Wacziarg

**ABSTRACT**

What obstacles prevent the most productive technologies from spreading to less developed economies from the world's technological frontier? In this paper, we seek to shed light on this question by quantifying the geographic and human barriers to the transmission of technologies. We argue that the intergenerational transmission of human traits, particularly culturally transmitted traits, has led to divergence between populations over the course of history. In turn, this divergence has introduced barriers to the diffusion of technologies across societies. We provide measures of historical and genealogical distances between populations, and document how such distances, relative to the world's technological frontier, act as barriers to the diffusion of development and of specific innovations. We provide an interpretation of these results in the context of an emerging literature seeking to understand variation in economic development as the result of factors rooted deep in history.

Enrico Spolaore
Department of Economics
Tufts University
Braker Hall
8 Upper Campus Road
Medford, MA 02155
and NBER
enrico.spolaore@tufts.edu

Romain Wacziarg
Anderson School of Management at UCLA
C-510 Entrepreneurs Hall
110 Westwood Plaza
Los Angeles, CA 90095-1481
and NBER
wacziarg@ucla.edu

# 1  Introduction

Technological differences lie at the heart of differences in economic performance across countries. A large and growing literature on development accounting demonstrates that total factor productivity accounts for a sizeable fraction of cross-country differences in per capita income (Hall and Jones, 1999, Caselli, 2005, Hsieh and Klenow, 2010, among many others). The problem of low technological advancement in poor countries is not primarily one of lack of innovation, for technologies that could make these countries vastly richer exist and are used elsewhere in the world. A major problem, instead, is one of delayed technological adoption. That many countries are subject to large technological usage gaps is a well-documented phenomenon. However, the factors explaining delayed technological adoption are not well-understood. What prevents the most productive technologies, broadly understood, from spreading to less developed economies from the world's technological frontier? In this chapter, we seek to shed light on this question, by quantifying the geographic and human barriers to the transmission of technologies.

We adopt a long-term perspective. The fortunes of nations are notoriously persistent through time, and much of the variation in economic performance is rooted in deep history.[1] While there have been reversals of fortune at the level of countries, these reversals are much less prevalent when looking at the fortunes of populations rather than those of geographic locations.[2] Indeed, contributions by Putterman and Weil (2010), Comin, Easterly and Gong (2010) and Spolaore and Wacziarg (2009, 2012a, 2013) argue that the past history of populations is a much stronger predictor of current economic outcomes than the past history of given geographical locations. Thus, any explanation for the slow and unequal diffusion of frontier technologies must be able to account for the persistence of economic fortunes over the long run. In this chapter, we argue that the intergenerational transmission of human traits, particularly culturally transmitted traits, has led to divergence between populations over the course of history. In turn, this divergence has introduced barriers to the diffusion of technologies across societies. These barriers impede the flow of technologies in proportion to how genealogically distant populations are from each other.

---

[1] For instance, an important literature has explored the prehistoric origins of comparative development (Diamond, 1997; Olsson and Hibbs, 2005; Ashraf and Galor, 2011 and 2013a).

[2] See Acemoglu, Johnson and Robinson (2002) for the reversal of fortune at the level of geographic locations (for former colonies), and papers by Spolaore and Wacziarg (2013) and Chanda, Cook and Putterman (2013) showing that the reversal of fortune disappears when correcting for ancestry and expanding the sample beyond former colonies.

Our starting point is to develop a theoretical model capturing these ideas. This model proceeds in three phases. Firstly, we argue that genealogical separation across populations leads, on average, to differentiation along a wide range of traits transmitted from parents to children either biologically or culturally. Populations that are genealogically distant should therefore also be distant in terms of languages, norms, values, preferences, etc. - a set of traits we refer to as vertically transmitted traits or more simply as vertical traits. Secondly, we consider the onset of a major innovation, which could be interpreted as the Industrial Revolution, and argue that differences in vertical traits introduce barriers to the diffusion of this major innovation across societies and populations. Thus, cross-country differences in aggregate TFP or per capita income should be correlated with their genealogical distance. Finally, we extend the model to allow for innovations taking place over time, and innovation and imitation occurring endogenously. In this more general framework, usage lags in the adoption of specific technologies and consequently aggregate differences in economic development are correlated with average differences in vertical traits, and thus with genealogical distance.

We next turn to empirical evidence on these ideas. To measure the degree of relatedness between populations, we use genetic distance. Data on genetic distance was gathered by population geneticists specifically for the purpose of tracing genealogical linkages between world populations (Cavalli-Sforza et al., 1994). By sampling large numbers of individuals from different populations, these researchers obtained vectors of allele frequencies over a large set of genes, or *loci*. Measures of average differences between these vectors across any two populations provide a measure of genetic distance. The measure we rely on, known as $F_{ST}$ genetic distance, is the most widely used measure in the population genetics literature because it has properties that make it well-suited to study separation times between populations - precisely the concept we wish to capture. $F_{ST}$ genetic distance has been shown to correlate with other measures of cultural differences such as linguistic distance and differences in answers to questions from the World Values Survey (Spolaore and Wacziarg, 2009, Desmet, Le Breton, Ortuño-Ortín and Weber, 2011).

Emphatically, the purpose of our study is *not* to study any genetic characteristics that may confer any advantage in development. The genes used in our measures of genealogical distance purposedly do not capture any such traits. It is important to note that the genes chosen to compare populations and retrace their genealogies are *neutral* (Kimura, 1968). That is, their spread results from random factors and not from natural selection. For instance, neutral genes include those

coding for different blood types, characteristics that are known *not* to have conferred a particular advantage or disadvantage to individuals carrying them during human evolutionary history. The mutations that give rise to specific alleles of these genes arise and spread randomly. The neutral genes on which genetic distance is based thus do not capture traits that are important for fitness and survival. As a result, measures based on neutral genes are like a molecular clock: on average they provide an indication of separation times between populations. Therefore, genetic distance can be used as a summary statistics for all divergence in traits that are transmitted with variation from one generation to the next over the long run, including divergence in cultural traits. Our hypothesis is that, at a later stage, when such populations enter in contact with each other, differences in those traits create barriers to exchange, communication and imitation. These differences could indeed reflect traits that are mostly transmitted culturally and not biologically - such as languages, norms of behavior, values and preferences. In a nutshell, we hypothesize that genetic distance measured from neutral genes captures divergence in intergenerationally transmitted traits - including cultural traits - between populations. This divergence in turn impedes the flow of innovations.

We use these measures of genetic distance to test our model of technological diffusion. Our barriers model implies that the genetic distance measured relative to the world technological frontier should trump absolute genetic distance as an explanation for bilateral income differences. We find this to be the case empirically. Our model also implies that genetic distance relative to the frontier should have predictive power for income differences across time even in periods when the world distribution of income was quite different from today's. We show indeed that the effect of genetic distance remains strong in historical data on population density and per capita income. Our model implies that after a major innovation, such as the Industrial Revolution, the effect of genealogical distance should be pronounced, but that it should decline as more and more societies adopt the frontier's innovation. This too is true empirically. Finally, our model implies that genetic distance should have predictive power at the level of disaggregated technologies, and find this to be the case both historically (when measuring technological usage on the extensive margin) and for more recent technological developments (measuring technological usage along the intensive margin). In sum, we find considerable evidence that barriers introduced by historical separation between populations are central to account for the world distribution of income.

In the final section of this chapter, we broaden our focus and place these hypotheses and findings in the context of the wider emerging literature on the deep historical roots of economic

development. Our discussion starts from a taxonomy, based on Spolaore and Wacziarg (2013), describing how historically transmitted traits could conceivably affect socio-economic outcomes. The taxonomy distinguishes between the *mode of transmission* of vertical traits, and the *mode of operation* of these traits. In principle, intergenerationally transmitted traits could be transmitted either biologically or culturally. However, the recent development of the literatures on epigenetics and on gene-culture interactions has made this distinction based on the mode of transmission much less clear-cut empirically and conceptually. A more fruitful discussion, we argue, is to try to better distinguish between the modes of operation of vertical traits. These traits, in principle, could bear direct effects on economic outcomes, or operate as barriers to economic interactions between populations. We discuss existing contributions in light of this distinction, and discuss directions for future research in the emerging new field concerned with the deep historical roots of economic development.

This chapter is organized as follows. Section 2 presents a stylized model of the diffusion of technologies as function of differences in vertically transmitted traits across human populations, and ultimately as a function of the degree of genealogical relatedness between them. Section 3 presents our empirical methodology and data. Section 4 discusses a wide range of empirical results pertaining to contemporaneous and historical measures of economic development and specific technology use measures. Section 5 discusses the interpretation of these results in the context of the broader literature on the deep roots of economic development. Section 6 concludes.

## 2    A Theory of Relatedness and Growth

In this section we present a basic theoretical framework to capture the links among genetic distance, intergenerationally-transmitted traits, and barriers to the diffusion of economic development across different societies.[3] The model illustrates two key ideas.

The first idea is that genetic distance between populations captures the degree of genealogical relatedness between populations over time, and can therefore be interpreted as a general metric for average differences in traits transmitted with variation across generations. Genetic distance measures the difference in gene distributions between two populations, where the genes under consideration are neutral. By definition, neutral genetic change tends to occur randomly, independently

---

[3] The model builds on Spolaore and Wacziarg (2009, 2012a).

of selection pressure, and regularly over time, as in a molecular clock (Kimura, 1968). This divergence provides information about lines of descent: populations that are closer in terms of genetic distance have shared a common "ancestor population" more recently. The concept is analogous to relatedness between individuals: two siblings are more closely related than two cousins because they share more recent common ancestors: their parents rather than their grandparents. Since a very large number of traits - not only biological but also cultural - are transmitted from one generation to the next over the long run, genetic distance provides a comprehensive measure for average differences in traits transmitted across generations. We call vertically transmitted traits (or vertical traits, for short) the set of characteristics passed on across generations within a population over the very long run - that is, over the time horizon along which populations have diverged (thousands of years).[4] Vertical transmission takes place across generations within a given population, and, in our definition, includes not only direct parent-to-child transmission of biological and cultural traits, but also, more broadly, "oblique" transmission of cultural traits from the older to the younger within a genetically-related group. In contrast, we define "horizontal transmission" as learning and imitation across different populations at a point in time.

The second idea is that differences in vertically transmitted traits act as barriers to horizontal learning and imitation, and therefore hamper the diffusion of innovations and economic development across societies.[5] We argue that populations that share a more recent common history, and are therefore closer in terms of vertical traits, face lower costs and obstacles to adopting each other's innovations. This view that differences in persistent societal characteristics may act as barriers is consistent with a large literature on the diffusion of innovations, starting with the classic work by Rogers (1962). Empirically, we are interested primarily in the diffusion of modern economic development in historical times, and especially after the Industrial Revolution, so our stylized model is designed with that objective in mind.

---

[4]This terminology is borrowed from the evolutionary literature on cultural transmission (for example, see Cavalli-Sforza and Feldman, 1981; Boyd and Richerson, 1985; Richerson and Boyd, 2005).

[5]Policy-induced barriers to the diffusion of technology are analyzed by Parente and Prescott (1994, 2002). In our framework we interpret barriers more broadly to include all long-term societal differences that are obstacles to the diffusion of development.

## 2.1 Genetic Distance and Vertically Transmitted Traits

We model all vertical traits of a population as a point on the real line: each population $i$ has vertical traits $v_i$, where $v_i$ is a real number. At time $o$ ("origin"), there exists only one population (population 0), with traits normalized to zero: $v_0 = 0$. At time $p > o$ ("prehistory"), the original population splits into two populations (1 and 2). At time $h > p$ ("history"), each of the two populations splits into three separate populations: population 1 into populations 1.1, 1.2, 1.3, and population 2 into populations 2.1, 2.2, and 2.3.[6] The genealogical tree is displayed in Figure 1. By analogy with the genealogy of individuals, we say that populations such as 1.1 and 1.2 are "sibling" populations, because their last common ancestors (their "parent" population) can be found at the more recent split (time $p$), while population pairs such as 1.2 and 2.1 are "cousin" populations, because their last common ancestors (their "grandparent" population) must be traced back to a more remote time $o < p$. $G(i, j)$ denotes the genetic distance between population $i$ and population $j$.[7] The genetic distance between two sibling populations is $g_s > 0$, while the genetic distance between two cousin populations is $g_c > g_s$. Formally,

$$G(1.m, 1.n) = G(2.m, 2.n) = g_s \quad \text{where } m = 1, 2, 3; \; n = 1, 2, 3 \text{ and } 1.m \neq 1.n \; ; \; 2.m \neq 2.n \quad (1)$$

and

$$G(1.m \; , \; 2.n) = g_c \quad \text{where } m = 1, 2, 3 \text{ and } n = 1, 2, 3 \quad (2)$$

Each population inherits vertical traits from its ancestor population with variation. In general, vertical traits $v_d$ of population $d$ (the "descendent"), descending from population $a$ (the "ancestor"), are given by:

$$v_d = v_a + \varepsilon_d \quad (3)$$

where $\varepsilon_d$ is a shock. In particular, we model the process of variation as a random walk. This simplification is consistent with the molecular-clock interpretation of genetic distance. While more complex processes could be considered, this formalization has two advantages: it is economical and illustrates how random changes are sufficient to generate our theoretical predictions. Formally, we assume that $\varepsilon_d$ takes value $\varepsilon > 0$ with probability $1/2$ and $-\varepsilon$ with probability $1/2$. We denote with

---

[6] In Spolaore and Wacziarg (2009), we presented a similar model with only four populations at time $h$ (1.1, 1.2, 2.1, and 2.2). Here we extend the framework to allow for a more general analysis, in which we also have pairs of populations that, while they are not at the frontier themselves, are both siblings with the frontier population.

[7] By defintion, $G(i, i) = 0$.

$V(i, j)$ the distance in vertically transmitted traits (vertical distance, for short) between populations $i$ and $j$:

$$V(i, j) \equiv |v_j - v_i| \tag{4}$$

We are now ready to summarize our first idea as:

**Proposition 1**

*The distance in vertical traits $V(i, j)$ between two populations $i$ and $j$ is, on average, increasing in their genetic distance $G(i, j)$.*

Derivation of Proposition 1:

The expected distance in vertical traits between sibling populations is:

$$E\{V(i, j) \mid G(i, j) = g_s\} = \varepsilon \tag{5}$$

because their vertical distance is equal to $2\varepsilon$ with probability $1/2$, when one population experiences a positive shock $\varepsilon$ and the other a negative shock $-\varepsilon$, and equal to 0 with probability $1/2$, when both populations experience the same shock (either $\varepsilon$ with probability $1/4$ or $-\varepsilon$ with probability $1/4$). In contrast, the expected distance in vertical traits between cousin populations is:

$$E\{V(i, j) \mid G(i, j) = g_c\} = \frac{3\varepsilon}{2} \tag{6}$$

because their vertical distance is 0 with probability $3/8$, $2\varepsilon$ with probability $1/2$, and $4\varepsilon$ with probability $1/8$.[8] Therefore, the expected distance in vertical traits is increasing in genetic distance:

$$E\{V(i, j) \mid G(i, j) = g_c\} - E\{V(i, j) \mid G(i, j) = g_s\} = \frac{\varepsilon}{2} > 0 \tag{7}$$

---

[8] The details of the calculation are as follows. With probability $1/4$, the two populations experienced identical shocks at time $h$, and their respective ancestor populations experienced identical shocks at time $p$, implying $V(i, j) = 0$. With probability $1/8$, one population lineage experienced a positive shock $\varepsilon$ at time $p$ and a negative shock $-\varepsilon$ at time $h$ while the other population lineage experienced $-\varepsilon$ and $\varepsilon$, implying again $V(i, j) = 0$. With probability $1/4$, the two populations' ancestors experienced identical shocks at time $p$, but the two populations experienced different shocks at time $h$, implying $V(i, j) = 2\varepsilon$. With probability $1/4$, the shocks were the same at time $h$ but different at time $p$, also implying $V(i, J) = 2\varepsilon$. Finally, with probability $1/8$, one population linaeage experienced two positive shocks ($\varepsilon + \varepsilon = 2\varepsilon$) and the other two negative shocks ($-\varepsilon - \varepsilon = -2\varepsilon$), therefore leading to a vertical distance equal to $4\varepsilon$. In sum, their expected vertical distance is given by $E\{V(i, j)|G(i, j) = g_c\} = \frac{3}{8}0 + \frac{1}{2}2\varepsilon + \frac{1}{8}4\varepsilon = \frac{3\varepsilon}{2}$.

It is important to notice that the relation between distance in vertical traits and genetic distance is is not deterministic, but works on average. Some pairs of populations, while genealogically more distant, may end up with more similar vertical traits than two more closely related populations. However, that outcome is less likely to be observed than the opposite. On average, genetic distance and vertical distance go hand in hand.

## 2.2 Barriers to the Diffusion of Economic Development

Our second idea is that differences in vertical traits constitute barriers to the spread of innovations across populations. A stylized illustration of this idea is provided below.

At time $p$ all populations produce output using the basic technology $Y_i = AL_i$, so that all populations have the same income per capita $y = A$. In period $h$ a population happens to find a more productive technology $A' = A + \Delta$ where $\Delta > 0$. We abstract from the possibility that the likelihood of finding the innovation is itself a function of a society's vertical traits. Such direct effects of vertical traits could strengthen the links between genetic distance and economic outcomes, but are not necessary for our results.

We denote the innovating population as $f$ (for technological frontier). To fix ideas and without loss of generality, in the rest of the analysis we assume that population 1.1 is the frontier population ($f = 1.1$). Populations farther from population $f$ in terms of vertical traits face higher barriers to adopt the new technology. Formally, we assume that a society $i$ at a vertical distance from the frontier equal to $V(i, f)$ can improve its technology only by:

$$\Delta_i = [1 - \beta V(i, f)]\Delta \tag{8}$$

where the parameter $\beta > 0$ captures the barriers to the horizontal diffusion of innovations due to distance in vertical traits. To ensure non-negativity, we assume that $\beta \leq \dfrac{1}{\max V(i, f)} = \dfrac{1}{4\varepsilon}$.[9] Therefore, income per capita in society $i$ will be given by:

$$y_i = A + \Delta_i = A + [1 - \beta V(i, f)]\Delta \tag{9}$$

This immediately implies:

### Proposition 2

---

[9] Alternatively, the formula could be re-written as $\Delta_i = \max\{[1 - \beta V(i, f)]\Delta, 0\}$.

*The difference in income per capita $|y_i - y_j|$ between society $i$ and society $j$ is a function of their relative vertical distance from the frontier $|V(i, f) - V(j, f)|$:*

$$|y_j - y_i| = \beta\Delta|V(i, f) - V(j, f)| \tag{10}$$

## 2.3  Genetic Distance and Income Differences

Since income differences are associated with differences in vertical traits across populations (Proposition 2), and differences in vertical traits, on average, go hand in hand with genetic distance (Proposition 1), we can now establish a link between expected income differences and genetic distance. These links are formally derived as Propositions 3 and 4 below.

### Proposition 3

*The expected income difference $E\{|y_j - y_i|\}$ between societies $i$ and $j$ is increasing in their genetic distance $G(i, j)$.*

Derivation of Proposition 3:

First, we must calculate the expected income of all pairs of populations at genetic distance $g_s$ (sibling populations). $V(i, j)$ between two sibling populations is 0 with probability $1/2$ and $2\varepsilon$ with probability $1/2$. When the two populations have identical traits, they have identical incomes. When they are at a distance $2\varepsilon$ from each other, one of them must be closer to the frontier's traits by a distance equal to $2\varepsilon$, no matter where the frontier's traits are located (at 0, $2\varepsilon$, or $-2\varepsilon$), or whether one of the two sibling populations *is* the frontier. Thus, when $V(i, j) = 2\varepsilon$, the income difference between the two populations is $\beta\Delta2\varepsilon$. In sum, for all pairs of sibling populations is $|y_{k.m} - y_{k.n}| = 0$ with probability $1/2$, and $|y_{k.m} - y_{k.n}| = \beta\Delta2\varepsilon$ with probability $1/2$, implying $E\{|y_{k.m} - y_{k.n}|\} = \beta\Delta\varepsilon$ where $k = 1, 2$; $m = 1, 2, 3$; $n = 1, 2, 3$; and $m \neq n$. Consequently, the expected income difference between sibling populations is:

$$E\{|y_j - y_i| \;\; || \; G(i, j) = g_s\} = \beta\Delta\varepsilon \tag{11}$$

Now, we must calculate the expected income difference between cousin populations. $V(i, j)$ between two cousin populations is 0 with probability $3/8$, $2\varepsilon$ with probability $1/2$, and $4\varepsilon$ with probability $1/8$. The calculation is slightly more complicated, because we must distinguish between pairs that include the frontier and pairs that do not include the frontier $f = 1.1$. First, consider pairs that

9

include the frontier. With probability $3/8$ a population $2.n$ shares the same traits (and hence income) with the frontier, with probability $1/2$, population $2.n$ has income lower than the frontier's by $\beta\Delta 2\varepsilon$, and with probability $1/8$ population $2.n$'s income is lower by $\beta\Delta 4\varepsilon$. Thus, we have:

$$E\{|y_f - y_{2.n}|\} = \frac{\beta\Delta 2\varepsilon}{2} + \frac{\beta\Delta 4\varepsilon}{8} = \frac{3\beta\Delta\varepsilon}{2} \qquad \text{where} \ \ n = 1,2,3 \qquad (12)$$

Now, consider pairs of cousin populations that do not include the frontier population - that is, pairs $1.m$ and $2.n$, with $m = 2,3$, and $n = 1,2,3$. Again, the income difference between each pair of cousin populations is equal to zero when both populations share the same traits (which happens with probability $3/8$), and is equal to $\beta\Delta 2\varepsilon$ when their traits are at a distance $2\varepsilon$ from each other (which happens with probability $1/2$), no matter where the frontier is located. However, when the two cousin populations are at a distance $4\varepsilon$ from each other (which happens with probability $1/8$), their income distance depends on the location of the traits of the frontier. If the frontier is at an extreme (either $2\varepsilon$ or $-2\varepsilon-$ an event with probability $1/2$), the $4\varepsilon$ vertical distance between $1.m$ and $2.n$ implies that their income distance is equal to $\beta\Delta 4\varepsilon$. In contrast, if the frontier's traits are at $0$ (also an event with probability $1/2$), $1.m$ and $2.n$ are equally distant from the frontier (each at a distance $2\varepsilon$), and therefore have identical incomes per capita. In sum, we have:

$$E\{|y_{1.m} - y_{2.n}|\} = \frac{\beta\Delta 2\varepsilon}{2} + \frac{1}{2}\frac{\beta\Delta 4\varepsilon}{8} = \frac{5\beta\Delta\varepsilon}{4}\text{where } m = 2,3; \ n = 1,2,3 \qquad (13)$$

Consequently, expected income difference between pairs of cousin populations is:

$$E\{|y_j - y_i| \ \ || \ G(i,j) = g_c\} = \frac{1}{9}\sum_{m=1}^{3}\sum_{n=1}^{3} E\{|y_{1.m} - y_{2.n}|\} = \frac{1}{9}[3\frac{3\beta\Delta\varepsilon}{2} + 6\frac{5\beta\Delta\varepsilon}{4}] = \frac{4\beta\Delta\varepsilon}{3} \qquad (14)$$

Therefore, the expected income difference between cousin populations is higher than the one between sibling populations: higher genetic distance is associated, on average, with higher income differences, as stated in Proposition 3. Formally:

$$E\{|y_j - y_i| \ \ || \ G(i,j) = g_c\} - E\{|y_j - y_i| \ \ || \ G(i,j) = g_s\} = \frac{\beta\Delta\varepsilon}{3} > 0 \qquad (15)$$

Why do populations which are genetically more distant from each other tend to differ more in income per capita, on average? The reason is that populations which are distant from each other genetically are also more likely to find themselves at more different distances from the frontier. Relative distance from the frontier, rather than genetic distance between populations per se, is the key determinant of expected income differences. Therefore, we can find an even stronger relation between income differences and genetic distance if we consider not the *absolute* genetic distance

between two populations $G(i, j)$, but their *relative* genetic distance from the technological frontier, defined as follows:

$$R(i, j) \equiv |G(i, f) - G(j, f)| \tag{16}$$

Our model predicts that the effect of relative genetic distance on income differences is not only positive, but also larger than the effect of absolute genetic distance:

**Proposition 4**

*The expected income difference $E\{|y_j - y_i|\}$ between society $i$ and $j$ is increasing in the two populations' relative genetic distance from the frontier $R(i, j)$. The effect of relative genetic distance $R(i, j)$ on income differences is larger than the effect of absolute genetic distance $G(i, j)$.*

Derivation of Proposition 4:

The expected income difference between pairs of populations at relative genetic distance $R(i, j) = g_s$ is[10]:

$$E\{|y_j - y_i| \ \| \ R(i, j) = g_s\}| = E\{|y_f - y_{1.2}|\} + E\{|y_f - y_{1.3}|\} = \beta \Delta \varepsilon \tag{17}$$

while the expected income difference between pairs of populations at relative genetic distance $R(i, j) = g_c$ is[11]:

$$E\{|y_j - y_i| \ \| R(i, j) = g_c\}| = \frac{1}{3} \sum_{n=1}^{3} E\{|y_f - y_{2.n}|\} = \frac{3\beta \Delta \varepsilon}{2} \tag{18}$$

Therefore, the effect of an increase of relative genetic distance from $g_s$ to $g_c$ is

$$E\{|y_j - y_i| \ \| \ R(i, j) = g_c\} - E\{|y_j - y_i| \ \| \ R(i, j) = g_s\} = \frac{\beta \Delta \varepsilon}{2} > \frac{\beta \Delta \varepsilon}{3} > 0 \tag{19}$$

The effect is positive ($\frac{\beta \Delta \varepsilon}{2} > 0$), and larger than the analogous effect of absolute genetic distance ($\frac{\beta \Delta \varepsilon}{3}$), derived above.

By the same token, the effect of relative genetic distance on expected income differences is also positive when moving from $R(i, j) = g_c - g_s$ to $R(i, j) = g_c$:

$$E\{|y_j - y_i| \ \| \ R(i, j) = g_c\} - E\{|y_j - y_i| \ \| \ R(i, j) = g_c - g_s\} = \frac{3\beta \Delta \varepsilon}{2} - \frac{5\beta \Delta \varepsilon}{4} = \frac{\beta \Delta \varepsilon}{4} > 0 \tag{20}$$

---

[10] We use the result, derived above, that all expected income differences between siblings are equal to $\beta \Delta \varepsilon$.

[11] We use the result, derived above, that the expected income difference between the frontier and each of its cousing populations is $\frac{3\beta \Delta \varepsilon}{2}$.

The results above are intuitive. As we increase relative genetic distance from the frontier, the expected income gap increases. The size of the effect is a positive function of the extent of divergence in vertically transmitted traits ($\varepsilon$), the extent to which this divergence constitutes a barrier to the horizontal diffusion of innovations ($\beta$), and the size of the improvement in productivity at the frontier ($\Delta$).

In summary, our model has the following testable implications, which are brought to the data in the empirical analysis carried in the rest of this chapter:

*1. Relative genetic distance from the frontier population is positively correlated with differences in income per capita.*

*2. The effect on income differences associated with relative genetic distance from the frontier population is larger than the effect associated with absolute genetic distance.*

## 2.4    A Dynamic Extension

In the stylized model above, for simplicity we assumed that only one big innovation took place at time $h$. We now present a dynamic example, where innovations take place over time, and innovation and imitation are modeled endogenously.[12] The key insights and results carry over to this extension.

In this dynamic example we assume, for simplicity, that populations do not change in modern times and have fixed size (normalized to one). More importantly, we assume that their inherited vertical traits do not change over the relevant time horizon. This is a reasonable simplification, because changes in vertical traits tend to take place much more slowly and at a longer horizon than the spread of technological innovations, especially if we focus on modern economic growth. Adding small random shocks to vertical traits after time $h$ would significantly complicate the algebra, but would not affect the basic results.

Consider our six populations ($i = 1.1, 1.2, 1.3, 2.1, 2.2, 2.3$), with vertical traits inherited from their ancestral populations as described above, and unchanged in modern times (i.e., for $t \geq h$). Time is continuous. Consumers in economy $i$ at time $t$ maximize:

$$U_i(t) = \int_s^\infty \ln c_i(s) e^{-\rho(t-s)} ds \tag{21}$$

under a standard budget constraint, where $c_i(t)$ is consumption, and $\rho > 0$ is the subjective discount rate. We assume that the six economies are not financially integrated, and that each economy $i$

---

[12]The model builds heavily on Barro and Sala-i-Martin (1997, 2003) and Spolaore and Wacziarg (2012a).

has its own real interest rate, denoted by $r_i(t)$. Hence, the optimal growth rate of consumption in society $i$ is:

$$\frac{dc_i}{dt}\frac{1}{c_i(t)} = r_i(t) - \rho \qquad (22)$$

The production function for final output $y_i(t)$ is:

$$y_i(t) = \int_0^{A_i(t)} [x_{zi}(t)]^\alpha dz, \quad 0 < \alpha < 1 \qquad (23)$$

where $x_{zi}(t)$ is the quantity of intermediate good of type $z$ employed at time $t$ in economy $i$, and the interval $[0, A_i(t)]$ measures the continuum of intermediate goods available in economy $i$ at time $t$. Each intermediate good is produced by a local monopolist.

As before, without loss of generality we assume that society 1.1 is the technological frontier ($f = 1.1$). In this setting, this means that $A_f(h) > A_i(h)$ for all $i \neq f$. However, unlike in the previous analysis, innovation at the frontier economy now takes place endogenously. Following Barro and Sala–i-Martin (1997 and 2003, chapters 6 and 8), we assume that the inventor of input $z$ retains perpetual monopoly power over its production within the frontier economy. The inventor sells the intermediate good at price $P_z = 1/\alpha$, earning the profit flow $\pi = (1 - \alpha)\alpha^{(1+\alpha)/(1-\alpha)}$ at each time $t$.

The cost of inventing a new intermediate good at the frontier is $\eta$ units of final output. Free entry into the innovation sector implies that the real interest rate $r_f(t)$ must be equal to $\pi/\eta$, which is assumed to be larger than $\rho$, therefore implying that consumption grows at the constant rate:

$$\gamma \equiv \frac{\pi}{\eta} - \rho > 0 \qquad (24)$$

Output $y_f(t)$ and the frontier level of intermediate goods $A_f(t)$ will also grow at the rate $\gamma$.

The other populations cannot use the intermediate goods invented in economy $f$ directly, but, as in Barro and Sala-i-Martin (1997), must pay an imitation cost $\mu_i$ in order to adapt those intermediate goods to local conditions. Our key assumption is that the imitation costs are *increasing in the distance in vertical traits between the imitator and the frontier*. Specifically, we assume that society $i$'s imitation cost is:

$$\mu_i(t) = \lambda e^{\theta V(i,f)} \left(\frac{A_i(t)}{A_f(t)}\right)^\xi \qquad (25)$$

This is an instance of our general idea: a higher $V(i, f)$ is associated with higher imitation costs, because differences in vertical traits between the imitator and the inventor act as barriers to adoption and imitation. The parameter $\theta$ captures the extent to which dissimilarity in vertical traits

13

between imitator and inventor increases imitation costs. For a given vertical distance, an imitator in society $i$ faces lower imitation costs when there is a larger set of intermediate goods available for imitation - that is, when $A_i(t)/A_f(t)$ is low. The rationale for this assumption is the usual one: intermediate goods that are easier to imitate are copied first. Hence, the parameter $\xi > 0$ captures this advantage from technological backwardness. Our perspective may indeed shed some light on whether backward economies face higher or lower imitation costs overall, an issue debated in the literature (for instance, see Fagerberg, 2004). As we will see, our model predicts that, in steady state, societies that are farther technologically, and should therefore face lower imitation costs for this reason (captured by the parameter $\xi$), are also farther in terms of vertical distance from the frontier, and hence should face higher imitation costs through this channel (captured by the parameter $\theta$), with conflicting effects on overall imitation costs.

Again, we assume that an imitator who pays cost $\mu_i(t)$ to imitate good $k$ has perpetual monopoly power over the production of that input in economy $i$, and charges $P_k = 1/\alpha$, earning the profit flow $\pi = (1 - \alpha)\alpha^{(1+\alpha)/(1-\alpha)}$, while output is proportional to available intermediate goods $A_i(t)$ in equilibrium: $y_i(t) = \alpha^{2\alpha/(1-\alpha)}A_i(t)$. With free entry into the imitation sector, economy $i$'s real interest rate in equilibrium is[13]:

$$r_i(t) = \frac{\pi}{\mu_i(t)} + \frac{d\mu_i}{dt}\frac{1}{\mu_i(t)} \tag{26}$$

In steady state, the level of imitation costs $\mu_i^*$ is constant. The number of intermediate goods, output and consumption in all economies grow at the same rate $\gamma$ as at the frontier. Therefore, in steady state the real interest rates in all economies are identical and equal to $\dfrac{\pi}{\eta}$, and imitation costs are identical for all imitators, which implies:

**Proposition 2bis**

*The difference in log of income per capita in steady state $|\ln y_i^* - \ln y_j^*|$ between society $i$ and society $j$ is a function of their relative vertical distance from the frontier $|V(i, f) - V(j, f)|$ :*[14]

$$|\ln y_i^* - \ln y_j^*| = \frac{\theta}{\xi}|V(i, f) - V(j, f)| \tag{27}$$

The intuition of the above equation is straightforward: long-term differences in total factor productivity and output between societies are an increasing function of their relative cost to imitate,

---

[13]See Barro and Sala-i-Martin (1997, 2003) for the details of the derivation.

[14]Of course we also have $|\ln A_i^*(t) - \ln A_j^*(t)| = |\ln y_i^*(t) - \ln y_j^*(t)|$

which depends on their relative vertical distance from the frontier. Therefore, societies that are more distant from the frontier in terms of vertically transmitted traits will have lower incomes per capita in steady state.

This dynamic model confirms the key implications of the simplified model that we had presented before. In particular, the equivalents of Propositions 3 and 4 hold in this setting as well, as long as one substitutes income differences $|y_j - y_i|$ with differences in log of income per capita in steady state $|\ln y_i^* - \ln y_j^*|$, and $\beta\Delta$ with $\dfrac{\theta}{\xi}$. We can then re-interpret those results as implying that societies at different relative genetic distance from the technological frontier will have different levels of income per capita in steady state. The effect of relative genetic distance on the income gap is larger when differences in vertical traits are associated with higher imitation costs (higher $\theta$). Interestingly, we also have that the effect of relative genetic distance on income differences is lower when there are larger benefits from technological backwardness (higher $\xi$). In sum, the effects of relative genetic distance on economic development extend to this dynamic setting.

# 3   Empirical Methodology and Data

## 3.1   Specification and Estimation

The starting points for our empirical investigation into the long-term barriers to economic development are Propositions 3 and 4. These theoretical results show that if differences in vertical traits act as barriers to the diffusion of technologies, then differences in measures of development or technological sophistication across pairs of countries should 1) be correlated with the absolute genetic distance between these countries, 2) be correlated more strongly with their genetic distance *relative to the technological frontier* and 3) genetic distance relative to the frontier should trump simple genetic distance between two countries. Whether these patterns hold true constitutes an empirical test of the barriers model. Denote by $D_i$ a measure of development or technological sophistiction in country $i$. We will consider alternatively per capita income (for the modern period), population density (for the pre-Industrial period) and direct measures of technology use, to be further detailed below. Denote by $FST_{ij}^W$ the absolute genetic distance between countries $i$ and $j$. Analogous to the theoretical definition, genetic distance *relative to the frontier* country is defined as: $FST_{ij}^R = |FST_{if}^W - FST_{jf}^W|$ where $f$ denotes the frontier country.

Then the empirical predictions of Propositions 3 and 4 lead to the following empirical specifi-

cations:

$$|D_i - D_j| = \alpha_0 + \alpha_1 FST_{ij}^R + \alpha_2' X_{ij} + \varepsilon_{ij}^\alpha \tag{28}$$

$$|D_i - D_j| = \beta_0 + \beta_1 FST_{ij}^W + \beta_2' X_{ij} + \varepsilon_{ij}^\beta \tag{29}$$

$$|D_i - D_j| = \gamma_0 + \gamma_1 FST_{ij}^R + \gamma_2 FST_{ij}^W + \gamma_3' X_{ij} + \varepsilon_{ij}^\gamma \tag{30}$$

where $X_{ij}$ is a vector of control variables, primarily composed of alternative sources of barriers to diffusion, primarily geographic barriers. The predictions of our model are that $\alpha_1 > 0$, $\beta_1 > 0$, $\alpha_1 > \beta_1$, $\gamma_1 > 0$ and $\gamma_2 = 0$.

Equations (28), (29) and (30) are estimated using least squares. However, an econometric concern arises from the construction of the left-hand side variable as the difference in development or technological sophistication across country pairs. Indeed, consider pairs $(i, j)$ and $(i, k)$. By construction, the log per capita income of country $i$ appears in the difference in log per capita incomes of both pairs, introducing some spatial correlation in the error term. To deal with this issue, we correct the standard errors using two-way clustering, developed by Cameron, Gelbach and Miller (2006). Specifically, standard errors are clustered at the level of country 1 and country 2. This results in larger standard errors compared to no clustering.[15]

We complement these tests with additional empirical results that can shed light on our barriers interpretation of the effect of genetic distance. In particular, we examine the evolution of the effect of genetic distance through time. If genetic distance continues to have an effect on differencs in economic performance in periods where the world distribution of income was very different, it should put to rest the idea that vertically transmitted traits bear direct, unchanged effects on productivity. We therefore examine the effects of genetic distance on population density in the pre-industrial era, going as far back as year 1. In Malthusian times, population density is the proper measure of overall technological sophistication, since per capita income gains resulting from innovation are only transitory, and soon dissipated by an increase in fertility (Ashraf and Galor, 2011 provide empirical evidence on this point). We also study the time path of the effect of genetic distance around the Industrial Revolution. Our model predicts that this effect should peak during the initial phases of the diffusion of the Industrial Revolution, as only places close to its birthplace

---

[15] In past work, we employed various methods to deal with the spatial correlation that arises as a byproduct of the construction of the left-hand side variable, such as including a set of common country dummies. The results were not sensitive to the method used to control for spatial correlation. See Spolaore and Wacziarg (2009) for further details.

have adopted the new innovation. The model predicts that the effect should decline thereafter, as more and more societies adopt industrial and post-industrial modes of production.

## 3.2 Data

### 3.2.1 Genetic distance data

Our source for genetic distance data is Cavalli Sforza et al. (1994). The main dataset covers 42 ethnolinguistic groups samples across the globe.[16] The genetic data concerns 120 gene loci, for which allele frequencies were obtained by population. The gene loci were chosen to represent neutral genes, i.e. genes that did not spread through natural selection but through random drift, as determined by geneticists. Thus, when aggregated over many genes, measures of genetic distance obtained from neutral genes capture separation times between populations, precisely the analog of genealogical distance employed in our theoretical model.

The specific measure of genetic distance we use is known as $F_{ST}$ genetic distance, also known as Wright's fixation index.[17] To illustrate the index, we derive it for the specific case of two populations, one *locus* and two alleles. The number of individuals in population $i$ is $n_i$. Total population is $n = \sum_{i=1}^{2} n_i$. The share of population $i$ is $w_i = n_i/n$. Consider one locus with two possible alleles: either $Q$ or $q$. Let $0 \leq p_i \leq 1$ be the frequency of individuals in population $i$ with allele $Q$. Let $p$ be this frequency in the whole population $\left(p = \sum_{i=1}^{2} w_i p_i\right)$. The degree of heterozygosity (i.e. the probability that two randomly selected alleles within a population are different) within population $i$ is $H_i = 2p_i(1 - p_i)$, and average heterozygosity across populations is $H_S = \sum_{i=1}^{2} w_i H_i$. Heterozygosity for the whole population is $H_T = 2p(1 - p)$. Then Wright's fixation index, $F_{ST}$, is defined as:

$$F_{ST} = 1 - \frac{H_S}{H_T} = 1 - \frac{n_1 p_1(1 - p_1) + n_2 p_2(1 - p_2)}{np(1 - p)} \tag{31}$$

This is one minus the ratio of group level average heterozygosity to total heterozygosity. If both populations have the same allele frequencies ($p_1 = p_2$), then $H_i = H_S = H_T$, and $F_{ST} = 0$. In the polar opposite case, individuals within each population all have the same alleles, and these

---

[16] We will also make use of a more detailed dataset covering 26 European populations. Since populations were sampled at the country level rather than at the ethnic group level for the European dataset, matching populations to countries was an easier task.

[17] In past work, we also used the Nei index. Results did not hinge on the use of either index.

alleles differ completely across groups ($p_1 = 1 - p_2$). Then $F_{ST} = 1$ (total fixation). In general, the higher the differences in allele frequencies across populations, the higher is $F_{ST}$. The formula can easily be extended to account for more than two alleles. $F_{ST}$ can be averaged in a variety of ways across loci, so that the resulting $F_{ST}$ distance is a summary measure of relatedness between the two populations. Moreover, boostrapping techniques can be used to obtain standard errors on estimates of $F_{ST}$. Details of these extensions are provided in Cavalli-Sforza et al. (1994, pp. 26-27). We rely on the genetic distance data that they provide, i.e. we rely on population geneticists' best judgment as to the proper choice of alleles, the proper sampling methods, and the proper way to aggregate heterozygosity across alleles.

The genealogical tree of human populations is displayed in Figure 2, where the genetic distance data was used to construct a tree showing the successive splits between human populations over the course of the last $70,000$ years or so. In this figure, recent splits indicate a low genetic distance between the corresponding populations. In the source data pertaining to 42 world populations, the largest $F_{ST}$ genetic distance between any two populations is between the Mbuti Pygmies and the Papua New Guineans ($F_{ST} = 0.4573$). The smallest is between the Danish and the English ($F_{ST} = 0.0021$).

Genetic distance is obtained at the level of populations but it was necessary to construct measures pertaining to countries. We matched ethnolinguistic groups in Cavalli-Sforza et al. (1994) to ethnic groups for each country using the ethnic group data from Alesina et al. (2003), and then constructed the expected distance between two individuals, each drawn randomly from each of the two countries in a pair. Thus, our baseline measure of genetic distance between countries 1 and 2 is:

$$FST_{12}^W = \sum_{i=1}^{I}\sum_{j=1}^{J}(s_{1i} \times s_{2j} \times FST_{ij}) \tag{32}$$

where $s_{1i}$ is the share of population $i$ in country 1, $s_{2j}$ is the share of population $j$ in country 2, and $FST_{ij}$ is genetic distance between population $i$ and $j$. This index is also known as the Greenberg index (after Greenberg, 1956), and is increasingly used in economics as a measure of ethnolinguistic heterogeneity (see for instance Bossert, D'Ambrosio and La Ferrara, 2011).[18]

The measure derived above, $FST_{12}^W$, is the absolute measure of expected distance between any

---

[18]In past work we also used the genetic distance between the largest populations (i.e. genetic groups) in countries 1 and 2. The correlation between expected (weighted) genetic distance and this alternative index is very high, and it does not matter which one we use in our empirical work.

two countries 1 and 2. In keeping with the theoretical definition, we can also define a measure of these countries' relative distance to the technological frontier $f$:

$$FST_{12}^{R} = |FST_{1f}^{W} - FST_{2f}^{W}| \tag{33}$$

Finally, the procedure above matches populations to ethnolinguistic groups as they occur in the contemporary period. It is, however, also possible to calculate genetic distance as of the year 1500 AD, by matching populations to the plurality group in each country given their composition in 1500. Thus, for instance, in the 1500 match, Australia is matched to the Aborigenes population (while for the contemporary period Australia is matched to a combination of English and Aborigenes - predominantly the former). We make use of the 1500 match in some historical regressions, or as an instrument for contemporary genetic distance. Again, measures of absolute and relative genetic distance are computed using the 1500 match of populations to countries.

### 3.2.2 Measures of development and technological sophistication

We use a variety of measures of differences in economic development and technological sophistication. The first set of measures is defined at an aggregate level. The primary measure for the contemporary period is the absolute difference in log per capita income in 2005 (from the Penn World Tables version 6.3). For the pre-industrial periods, we consider the absolute difference in population density. The population density data pertains to the year 1500, and the source is McEvedy and Jones (1978). Despite more limited geographic coverage, we also use data on per capita income going back to 1820, from Maddison (2003), in order to examine the time path of the effect of genetic distance around the time of the Industrial Revolution.

The second set of measures includes disaggregated measures of technology usage, either along the extensive margin (for the historical period) or along the intensive margin (for the contemporary period).[19] We rely mostly on data from Easterly, Comin and Gong (2010, henceforth CEG). CEG gathered data on the degree of technological sophistication for the years 1000 BC, 1 AD, 1500 AD and the contemporary period (1970-2000 AD). We make use of the data for 1500 AD and the contemporary period, since this corresponds most closely to the available genetic distance data. The data for 1500 pertain to the extensive margin of adoption of 24 separate technologies, grouped into 5 categories: military, agricultural, transportation, communication and industry. For each

---

[19]These technologies are listed in Appendix 1.

technology in each category, a country is given a score of 1 if the technology was in use in 1500, 0 otherwise. The scores are summed within categories, and rescaled to vary between 0 and 1. An overall index of technological sophistication is also obtained by taking the simple average of the technological index for each of the 5 categories.

For the 1970-2000 AD data, technology usage is measured along the intensive margin. The basic data covers the per capita usage intensity of nine technologies, obtained from the database of Comin, Hobijn and Rovito (2008). For each technology, a country's usage is characterized as the number of years since the technological frontier (the United States) had the same level of per capita usage. The index is then rescaled to vary from 0 to 1, where 1 denotes usage at the same level as the frontier. Technologies are aggregated into 4 of the 5 aforementioned categories (all except the military category), and a simple average of the four measures is also available.

Finally, we attempted to measure technological sophistication at a more disaggregated level. This allows for a more refined analysis based on individual technologies that were not aggregated into broader categories, as is the case in the CEG dataset. For this, we relied on the CHAT dataset (Comin and Hobijn, 2009), which contains data on usage of 100 technologies. We restricted attention to technologies for which data is available for at least 50 countries over the 1990-1999 period. This led to a restricted set of 33 technologies, covering a wide range of sectors - medical, transportation, communications, industrial and agricultural technologies. For each of the underlying 33 technologies, we calculated usage per capita, in order to maintain a consistent definition of the intensity of use.[20] For instance, for the technology "personal computers", the dependent variable is the absolute difference, between country $i$ and country $j$, in the number of computers per capita. For all technologies, the technological leader was assumed to be the United States, an assumption confirmed in virtually all cases when examining the actual intensity of use.

All of these measures of technological sophistication were available at the country level, so we computed the absolute difference in technology measures across all available pairs of countries for the purpose of empirical analysis.

---

[20]One exception was for the share of cropland area planted with modern varieties, for which it would make little sense to divide by population. All other technologies were entered in per capita terms.

### 3.2.3 Measures of geographic barriers

Measures of genetic distance are correlated with geographic distance. Indeed, *Homo sapiens* is estimated to have migrated out of East Africa around 70,000 years ago, and from there spread first to Asia, and then later fanned out to Europe, Oceania, and the Americas. As early humans split into subgroups, the molecular clock of genetic drift started operating, and populations became more genetically distant. It is not surprising that the farther in space, the more genetically distant populations are expected to be. It is therefore important to control for geographic distance when estimating the human barriers to the diffusion of innovations. At the same time, as we describe below, the correlation between geographic distance and genetic distance is not as large as one might expect. This is the case for two major reasons: First, genetic drift occurred along rather specific geographic axes. For instance, a major dimension along which populations array themselves in proportion to their genetic distance is a rough straight line between Addis Ababa and Beijing. There need not be a strict correspondence, then, between genetic distance and common measures of geographic distance relevant as geographic barriers to the spread of innovations, such as the greater circle distance or latitudinal distance. Second, more recent population movements have served to break the initial links between geographic distance and genetic distance. Two highly relevant population movements were the conquests of parts of the New World by Europeans, and the slave trades occuring thereafter. We obtain some (but not all) of our identifying variation off of these post-1500 population movements.

To capture geographic distance we use a large array of controls, capturing both simple geodesic distance, distance along the longitudinal and latitudinal dimensions, and binary indicators of micro-geography such as whether the countries in a pair are contiguous, are islands, are landlocked, or share a common sea or ocean. This set of controls was included in every regression, and was supplemented in robustness tests by additional geographic controls such as climatic differences, continent effects and freight costs.

### 3.2.4 Summary statistics and data patterns

Figure 3 presents a simple plot of weighted genetic distance to the USA against per capita income, and Figure 4 does the same after partialling out the effect of geodesic distance (a similar figure obtains after partialling out the effect of a longer list of geographic distance metrics). Both figures reveal a negative association between per capita income and genetic distance to the USA.

Table 1 presents summary statistics to help in the interpretation of regression estimates. Panel B displays correlations based on $10,440$ country pairs, based on 145 countries. These correlations are informative: the absolute genetic distance between pairs bears a correlation of 19.5% with the absolute difference in log per capita income. Genetic distance relative to the USA, however, bears a much larger correlation of 32.26%, a pattern consistent with the predictions of the barriers model, implying a larger effect of relative genetic distance compared to absolute genetic distance. Finally, as mentioned above, the correlation between genetic distance (either relative to the frontier or not) with geodesic distance is positive but moderate in magnitude, offering hope that the effect of genealogical barriers can be estimated separately from that of geographic barriers.

# 4   Barriers to Development: Empirical Results

## 4.1   Results for Aggregate Measures of Economic Development

### 4.1.1   Baseline estimates

Baseline estimates of equations (28), (29) and (30) are presented in Table 2. The predictions of the barriers model are borne out: after controlling for various measures of geographic distance, differences in per capita income are significantly correlated with both absolute and relative genetic distance (columns 1 and 2).[21] However, the magnitude of the effect of genetic distance relative to the technological frontier (column 1) is about three times as large as the effect of absolute genetic distance (column 2). This is true when comparing both the estimated coefficient and a standardized measure of magnitude (the standardized beta, reported in the next to last row of Table 2). When including both measures in the regression (column 3), genetic distance relative to the frontier remains significant while absolute genetic distance becomes insignificantly different from zero. In terms of magnitudes, a one standard deviation increase in $F_{ST}$ genetic distance relative to the USA is associated with an increase in the absolute difference in log income per capita of almost 29% of that variable's standard deviation.

Column 4 of Table 2 reports results of IV estimation, using relative genetic distance to the English population in 1500 as an instrument for current genetic distance to the USA. This is meant to

---

[21] A myriad additional controls were included as robustness tests in analogous regressions presented in Spolaore and Wacziarg (2009). These included climatic differences, freight costs, etc. Results were robust to the inclusion of these additional control variables.

address two specific concerns: First, matching the 42 populations for which genetic distance data is available to contemporaneous ethnolinguistic groups may introduce measurement error. The main difficulties in the match arise for the New World where it is sometimes difficult to assess which European population to match with the descendents of past European settlers, which African populations to match with former slaves, and what shares to ascribe to these various populations in the total population, given that many of them mixed over time, resulting in significant shares of populations with mixed ancestry (the latter issue arises mainly in Latin America). In contrast, the 1500 match of genetic groups (populations) to the plurality ethnic group is much more straightforward, since the Cavalli-Sforza et al. (1994) data was gathered precisely to represent the makeup of countries as they stood in 1492, prior to the population movements associated with the conquest of the New World. The second concern is endogeneity: genetic distance between countries changed in the post-1492 era due to the aforementioned conquest of the New World and the slave trades. It is possible that areas well-suited to high incomes in the Industrial Era, perhaps due to geographic factors such as a temperate climate, happened to attract certain populations (for instance Europeans) as settlers. In this case, it would be the potential for differential incomes that would causally affect genetic distance rather than the opposite. Using genetic distance lagged by 500 years as an instrument addresses this particular endogeneity concern. The results presented in column 4 show that, if anything, OLS understated the effect of relative genetic distance: its standardized effect rises under IV to 46.49%. Since the IV estimates are larger than the OLS estimates, to remain conservative we rely in the rest of this chapter on OLS estimates.

### 4.1.2 Regional controls and analysis

In Table 3, we run a variety of regressions accounting for regional effects. In column 1, we include a full set of continental dummy variables capturing both whether the countries in a pair are both located on the same specific continent (an effect presumed to go in the direction of reducing the difference in economic performance between these countries) and whether they are located on different ones (as further defined in the footnote to Table 3). The idea behind this test is to further control for geographic factors not already captured by the included geographic distance variables. However, this is a demanding test, since continent effects could capture geographic barriers but also part of the effect of human barriers that could be mismeasured when using genetic distance. Nonetheless, the effect of genetic distance remains robust to controlling for a full set of twelve same- and different-continent dummies. While the effect of genetic distance falls in magnitude, it remains

large and highly significant statistically.

Columns 2 and 3 make use of the separate genetic distance dataset we have for 26 countries in Europe. Here, the relevant measure of genetic distance is $F_{ST}$ distance to the English (the birthplace of the Industrial Revolution), though the results do not change if we use distance to the Germans instead. We find that within Europe, genetic distance is again a strong predictor of absolute differences in log per capita income. The standardized beta on genetic distance relative to the English is of the same order of magnitude as that found in the world sample, and it is highly significant. There are two major genetic clines in Europe: one separating the North and the South, another one separating the East and the West. These correspond to North-South and East-West income differences. Since the East-West cline overlaps to a large degree with regions that were on either side of the Iron Curtain during the Cold War, to assess whether this historical feature explains all of the effect of genetic distance on economic performance we repeat our regression using income in 1870 (from Maddison), well prior to the rise of the Eastern bloc. We find that the effect of genetic distance is in fact larger in magnitude in the immediate aftermath of the Industrial Revolution, with the standardized beta rising to almost 44%. This results assuages concerns that the contemporary results were a result of the fact that the Iron Curtain as a first approximation separated Slavic from non-Slavic Europeans. It is also highly consistent with the barriers story since, as we further explore below, the effect of genetic distance should be larger around the time of a large innovation, in the midst of the process whereby countries other than the frontier are busy adopting the frontier technology in proportion to how genetically far they are from the frontier. In sum, our effects hold within Europe, where genetic distance is better measured.

Since the basic result of this chapter holds so strongly for Europe, might Europe drive the World results? To test this, in column 4 we exclude any pairs of countries containing at least one European country. Compared to the baseline results, the standardized effect of genetic distance relative to the USA declines from 30% to 25%, but remains large and statistically significant - highlighting that the results are not due to Europe alone. To drive home the point, in column 5 we control for the absolute difference in the share of the population of European descent, using data from the Putterman and Weil (2010) migration matrix. The regression now controls more broadly for the effect of Europeanness, and while the effect of the absolute difference in the share of Europeans is a positive and statistically significant determinant of differences in per capita income, its inclusion in the regression only moderately reduces the standardized effect of relative genetic distance (to

27%). We conclude that our results are not driven by the inclusion of European countries in the sample, nor are they driven by the genetic difference between Europeans and the rest.

The final geographic concern that we explore is whether Sub-Saharan Africa drives our results. As Figure 2 illustrates, Sub-Saharan African populations are genetically distant from the rest of the world: the Out of Africa migrations occurring about 70,000 years were the first foray of modern humans out of Africa, and consequently Africans and other world populations have had the longest time to drift apart genetically from each other. Sub-Saharan populations also have some of the lowest pre capita GDPs recorded in the world. While it is part of our story to ascribe some of the poverty of Africa to the barriers to technological transmission brought about by its high degree of genealogical distance from the rest of the world, it would be concerning if our results were entirely driven by Sub-Saharan Africa. To address this concern, in column (6) of Table 3 we exclude any pair that involves at least one Sub-Saharan country from our sample. We find that the effect of genetic distance falls a little, but remains positive, statistically significant, and large in magnitude with a standardized beta equal to 17%. Together with the strong results within Europe, this should lay to rest any notion that our results are driven solely by Sub-Saharan Africa.

### 4.1.3 Historical analysis

We now turn to a historical analysis of the determinants of aggregate measures of economic performance, seeking to achieve two main goals. The first is to assess the robustness of the effect of genetic distance through time. The second goal is to describe the time path of the standardized effect of genetic distance around the time of the Industrial Revolution. In our barriers model, a major innovation such as the Industrial Revolution should lead to a specific pattern in the evolution of the effect of relative genetic distance on differences in economic development. Specifically, the effect of genetic distance should be large in the aftermath of the onset of the Industrial Revolution in the frontier country. As more and more societies adopt the Industrial Revolution, the effect should gradually decline. We now redefine the frontier country as the United Kingdom (i.e. the English population) since it is a more appropriate choice for the period concerned.[22]

---

[22]This choice is not very material. In fact, relative genetic distance to the English and relative genetic distance to the United States are very highly correlated, because the United States are primarily composed of populations from Western Europe - either the English or populations genetically very close to the English. In fact, by world standards genetic distances betweem Western European populations are so small that it matters little empirically which Western European population is chosen as the frontier. For instance, for 1500 we experimented with using

Table 4 displays pairwise correlations between historical measures of differences in economic development and genetic distance. For the 1500 period, we consider the correlation between relative genetic distance to the English using the 1500 match, and population density. For periods from 1820 to today, it is best to rely on the correlation between contemporaneous weighted genetic distance relative to the UK, and the absolute difference in log per capita income at various dates.[23] A few remarks are in order: First, this data reveals some persistence in economic fortunes. In spite of being different measures, even the absolute difference in population density in 1500 and the absolute difference in log per capita income in 2005 bear a correlation of about 12% with each other. Correlations between income-based measures are much higher (for instance the correlation of income differences in 1820 and 2005 is 33%). Second, genetic distance is positively and significantly correlated with these measures of differences in economic performance at all dates. For instance, the correlation between the absolute difference in population density in 1500 and relative genetic distance to the English in 1500 is about 16%. This rises to 32% in 2005 (comparisons of magnitudes should be made cautiously from this table as the underlying samples differs by date - but in the case of 1500 and 2005 the samples are very similar - more on this point below). In general, simple correlations reveal that despite some changes in the relative fortunes of nations over the last 500 years, the correlation between genetic distance and development seems to exist at all dates.

Table 5 turns to regression analysis. Across all columns, corresponding to different dates, genetic distance relative to the UK comes out with a positive, significant coefficient. Thus, the effect of genetic distance is robust to considering different dates and a different measure of economic development for the Malthusian period. The penultimate row of Table 5 shows the evolution of the standardized effect of genetic distance over time for a common sample of 820 country pairs (41 countries), for which income data is available at all dates. The magnitudes here are somewhat smaller than when using unrestricted samples across periods, in part because the 41 countries only include one Sub-Saharan African country (and that country is South Africa, which is relatively rich). However, restricting the sample to pairs available at all dates allows for a comparison of magnitudes across time. To facilitate interpretation, the standardized effects from the common

---

Italy as the frontier country; results were unchanged.

[23]We lack genetic distance data suitable for the millenia prior to 1500, despite the existence of some population density data for early dates. At any rate it is not clear that our barriers story would apply with as much force in periods where geographic barriers to the diffusion of innovation were so overwhelming, except perhaps in a regionally narrow context.

sample are displayed in Figure 5.

This figure lends further credence to the barriers model. Indeed, just as predicted above, the effect of genetic distance, which is initially modest in 1820, rises by around 75% to reach a peak in 1913, and thereafter declines. Thus, in the few decades following the adoption of the Industrial Revolution by countries in the (genetic) periphery of England, the effect of genetic distance was maximal. Thereafter, as more and more societies industrialized, the effect fell steadily.

## 4.2 Results for specific innovations

The analysis above concerns determinants of differences in aggregate productivity. This is useful to analyze very broad trends like the diffusion of the Industrial Revolution. Yet our model also applies to the diffusion of more specific technologies. Indeed, if our empirical results applied to aggregate measures of development or technological sophistication only, but did not extend to more disaggregated technologies, it would cast doubt on the hypothesis that the main effect of genetic distance is to hinder the transmission of technologies across societies with very different cultures and histories. In this subsection, we use data directly at the technology usage level to address this issue.

Table 6 starts with some summary statistics from the CEG dataset, pertaining to the contemporary period. Panel A is mainly meant to assist in the interpretation of the regressions that come next, but Panel B already contains interesting information. The first observation is that differences in the intensity of technology usage in 1970-2000 across various technological categories are correlated, but imperfectly. Second, differences in technology usage intensity are positively correlated with per capita income, but the correlations are in the $0.4 - 0.7$ range depending on the technological category, so these variables do not all measure the same thing. In other words, our measures of differences in technology usage are not simply indicators of differences in overall economic performance. Third, differences in technology usage are correlated more strongly with genetic distance relative to the frontier than with genetic distance per se. In fact, correlations with the latter are often close to zero while correlations with the former are always positive and significant.

Table 7 carries out the regression analysis for the contemporary period, controlling for geographic distance. Genetic distance relative to the frontier comes out positive in all cases, and significant at the 5% level or better for 3 of the 4 technological categories, as well as for the summary

index of overall technology usage. The only category for which genetic distance is not significant is agricultural technologies. One possible interpretation is that agricultural technologies, for the contemporary period under consideration, have already widely diffused around the globe and are already intensively in use in much of the developing world, so that the effect of genetic distance as a barrier to their adoption can no longer be detected. We also carried out the same regression analysis as that in Table 8, but adding to the specification the measure of absolute genetic distance between pairs.[24] We found that relative genetic distance always trumped absolute distance, which sometimes carried a negative sign and was statistically insignificant in most cases. Thus, our test of the barriers story (equation 30) also works when considering technology usage intensity rather than aggregate measures of development.

Turning to the historical evidence, Table 8 examines the determinants of technology usage differences along the extensive margin in the year 1500. As before, we use the English population as the frontier (as before, it matters little if we use the Italians instead - Italy was arguably the most technologically sophisticated country in the world in 1500). For 1500 we have 5 rather than 4 technological categories, plus the overall index of technological sophistication. We find that in all cases, genetic distance relative to the English is positive and statistically significant at the 10% level. In 5 of the 6 columns, it is significant at the 1% level (as before, the weakest results are for agricultural technologies). This is remarkable given the crudeness of the measure of technological use in 1500, based on counting whether or not each of 24 technologies, grouped in functional categories, were in use at all in a given country at the time. Moreover, as before we also conducted horseraces between relative genetic distance and absolute genetic distance.[25] For five of the six indicators we again found that relative genetic distance trumps absolute genetic distance, with the latter entering with either the wrong sign, a very small magnitude, or low significance levels. The only exception, once again, was for agriculture.

Finally, we carried out the same analysis with the 33 disaggregated technologies chosen from the CHAT dataset. The results are presented in Table 9. For each technology, the table reports the coefficient on relative genetic distance to the USA (from a regression in which the standard set of geographic controls are included), the number of observations and countries, the standardized beta coefficient on genetic distance, and the $R^2$. The results vary across technologies of course, but

---

[24]Results are available upon request.

[25]Results are available upon request.

interesting observations emerge: 1) In *every single case* the effect of genetic distance on differences in technology usage intensity is positive. 2) In 22 of the 33 cases, the coefficient on genetic distance is significant at the 10% level, and in 19 cases at the 5% level. 3) The effect of genetic distance is particularly strong for disaggregated agricultural technologies and industrial technologies, and weakest for transportation and medical technologies. 4) The magnitude of the standardized effects, for those that are statistically significant, vary from 8% to 24%, a bit smaller but roughly in line with what we found using aggregate measured of productivity or the CEG dataset.[26]

A consideration of technologies at a more disaggregated data, rather than measures of overall productivity at the economy-wide level, provides additional evidence that human barriers matter. Not only is genetic distance relative to the frontier a strong predictor of technological usage differences in 1500 and in the contemporary period, we also find that it generally trumps absolute genetic distance. The fact that genetic distance accounts for differences in technological usage indicates that our previous aggregate results might in large part be accounted for by hindrances to the adoption of frontier technology brought about by historical separation between populations.

# 5    Ancestry and Long Run Development

In this section, we broaden the discussion of the role of ancestry as a determinant of the comparative wealth of nations, building on the discussion in Spolaore and Wacziarg (2013).[27] Our basic argument is that traits passed on across generations within societies play a fundamental role in accounting for the persistence of economic fortunes. However, the specific way in which these traits operate can take a variety of forms. In the model presented above, we argued that differences in vertically-transmitted traits introduced *barriers* to the diffusion of innovations across nations. We found much evidence that this was the case for aggregate productivity and for specific innovations going back to the year 1500. However, we have not said much about what causes the *onset* of these innovations. Other authors have pointed to a role for traits to bear a *direct* effect on the onset of major productivity enhancing innovations, broadly construed. We have also not said much

---

[26] We also conducted horseraces between absolute and relative genetic distance for each of the 33 disaggregated technologies. Relative genetic distance remains positive and significant in 17 of the 22 cases where relative genetic distance is significant at the 10% level when entered on its own. In the vast majority of cases, absolute genetic distance enters insignificantly or with a *negative* sign.

[27] The discussion of the relation between cultural traits and economic outcomes is also drawn in part from Spolaore (2014).

about the nature and specific method of transmission of the traits that are thought to matter for prosperity. These traits could be transmitted culturally, biologically, or through the interaction of culture and biology.

We proceed in several steps. We start by briefly describing the growing literature on long run persistence in the wealth of nations. We argue that the intergenerational transmission of traits has a lot to do with explaining long-run persistence, because traits are much more easily transmitted across generations that across societies. That is, ancestry matters to explain the wealth of nations. Next, we introduce a taxonomy to understand the manner in which ancestry matters. In particular, we introduce a distinction between barrier effects and direct effects of vertical traits. We also distinguish between the mode of transmission of the traits, either cultural, biological or dual. Finally, we provide several examples from the recent literature illustrating the various ways in which ancestry can matter.

## 5.1 Persistence and Reversals: The Role of Ancestry

Discussions of the long run roots of comparative development usually starts with geographic factors. A large literature has documented strong correlations between economic development and geographic factors, for instance latitude, climate and the disease environment.[28] The observation that geographic factors are correlated with development was at the root of Diamond's 1997 book on the long-run development advantage enjoyed by Eurasia - particularly Europe. On the surface, geography is a convenient explanation for persistence, because geography does not change very much, so that this immutable factor can be thought of as a prime reason for persistence in the wealth of nations. This view, however, is overly simplistic, for at least two reasons: First, the effect of geography on economic outcome can change depending on the technology of production. Geographic features useful to produce GDP in an agrarian economy may not be as helpful in an industrial society. Second, the manner in which geographic factors affect development today is open to a variety of interpretations. The factors could operate directly (for instance a high disease burden can reduce productivity) or have an indirect effect through their historical legacy. While both channels could be operative, the literature has increasingly moved in the latter direction.

---

[28]See, for instance: on climate and temperature, Myrdal (1968); Kamarck (1976); Masters and McMillan (2001); Sachs (2001). On the disease environment: Bloom and Sachs (1998); Sachs, Mellinger and Gallup (2001); Sachs and Malaney (2002). On natural resources: Sachs and Warner (2001).

In fact, Jared Diamond (1997) pointed out early that geographic factors such as the shape of continents and the availability of domesticable plants and animals probably did not have much to do with current development directly. It is because these factors gave people from Eurasia an early advantage in development, and because this advantage has persisted through the generations, that Europeans were able to conquer the New World (and many parts of the old one) and to remain at the top of the world distribution of income for a long time. This point became more widely recognized since a pathbreaking paper by Acemoglu, Johson and Robinson (2002) where these authors pointed out that the reversal of fortune experienced by former colonies between 1500 and today was inconsistent with a simple, direct effect of geography: for the geographic factors that made countries poor five hundred years ago should be expected to make them poor today still. And yet fortunes were reversed among a significant portion of the world's countries. This paper pointed to an indirect effect of geography, operating through institutions: where Europeans settled, they brought good institutions, and these are the fundamental proximate cause of development. Where Europeans chose to exploit and extract, the institutions they bequeathed had negative effects on development.

Yet that interpretation, too, became the subject of debates. Glaeser, La Porta, Lopez-de-Silanes and Shleifer (2004), for instance, state: "the Europeans who settled in the New World may have brought with them not so much their institutions, but themselves, that is, their human capital. This theoretical ambiguity is consistent with the empirical evidence " We would go even further: Europeans who settled in the New World brought with them the whole panoply of vertically transmitted traits - institutions, human capital, norms, values, preferences. This vector of vertical traits was by definition easier to transmit to the descendents of Europeans than it was to convey to colonized populations. This interpretation suggests an important role for ancestry, rather than only institutions, as an explanation for the reversal of fortunes. Locations that were colonized by Europeans and were previously characterized by low population density and the prevalence of non-agrarian modes of subsistence became rich. Locations that were inhospitable to Europeans remained poor, and Europeans remained at the top of the world distribution of aggregate productivity throughout.[29] That the wealth of a nation seems so strongly affected by the wealth of the

---

[29] We greatly expand on this point in Spolaore and Wacziarg (2013). In that paper, we revisit the Acemoglu, Johnson and Robinson (2002) evidence on the reversal of fortune. By examining the correlation between population density in 1500 and per capita income today, we confirm their findings for former colonies. Yet we also show that: 1) any evidence of a reversal of fortune disappears when European countries are included in the sample; 2) there

ancestors of those living in that nation suggests a central role for vertically transmitted traits as an explanation for both long-run persistence and the current distribution of income.

This interpretation led various authors to focus explicitly on persistence and ancestry. First came our own work on genetic distance as a barrier to development, already discussed in the previous sections (Spolaore and Wacziarg, 2009). Next came important papers by Putterman and Weil (2010) and Comin, Easterly and Gong (2010). These papers also explore the deep historical roots of current development.

Putterman and Weil (2010) look at two important determinants of the current wealth of nations: experience with agriculture, measured by the time elapsed since the adoption of sedentary agriculture as a primary means of food production; and experience with a centralized state, measured by the number of years a country has experienced centralized governance, discounting years that occurred farther in the past. Both variables are predictors of today's per capita income, but they enter even more strongly when they are adjusted for ancestry. To adjust variables for ancestry, Putterman and Weil construct a migration matrix. In this matrix, a row pertains to a country, and columns contain the fraction of that country's population whose ancestors in 1500 lived in each of the world's countries. For the Old World, entries are mostly diagonal: that is, the ancestors of the French mostly lived in France in 1500. For the New World, however, the ancestors of current populations are often in significant numbers from other continents altogether - primarily European countries for European colonizers, and Sub-Saharan African countries for the descendants of former slaves. By premultiplying a variable by the migration matrix, one obtains this variable's ancestry-adjusted counterpart. For instance, for Australia the history of the *location* is the history of the Aborigenes, while the history of the current *population* is mostly the history of the English. Putterman and Weil's major contribution is to show that ancestry-adjusted years of agriculture and ancestry-adjusted state centralization are much stronger predictors of current income than their non-ancestry adjusted counterparts. This suggests an important role, again, for traits that are passed on intergenerationally within populations.

Comin, Easterly and Gong (2010) take a different approach, but reach a similar conclusion: they show that the degree of technological sophistication of countries is highly autocorrelated even at

is evidence of persistence among countries that were not former European colonies; 3) persistence is even stronger when looking at countries that are populated mostly by their indigenous populations. These facts are suggestive of a strong role for ancestry as an explanation for persistence.

very long horizons: they detect correlations between current technological usage levels (measured along the intensive margin in the current period) and technological usage as far back as the year 1,000 BC (measured along the extensive margin for a set of 12 ancient technologies). Current per capita income is also correlated strongly with past technological sophistication in the years 1,000 BC, 1 AD and 1500 AD. In this case, a history of technological advancement predicts current income and technological advancement, an indication of persistence. The crucial point, however, is again that when the historical (lagged) variables are entered in their ancestry-adjusted forms, they are much stronger predictors of current outcomes than variables that capture the history of a location. In this context also, there appears to be a strong role for ancestry and intergenerational transmission as explanations for the persistence in technology and income levels.

Why does ancestry matter? In what follows we present a taxonomy of the possible effects of vertically transmitted traits on growth and development. This taxonomy is summarized in the following matrix:

| Mode of Operation ⟶ Mode of Transmission ↓ | Direct Effect | Barrier Effect |
|---|---|---|
| **Biological Transmission** (genetic and/or epigenetic) | Quadrant I | Quadrant IV |
| **Cultural Transmission** (behavioral and/or symbolic) | Quadrant II | Quadrant V |
| **Dual Transmission** (biological-cultural interaction) | Quadrant III | Quadrant VI |

## 5.2  Modes of Transmission

The inheritance of traits from one generation to the next in humans takes place through several modes of transmission and along multiple dimensions. Recent scientific advances stress the complexity of different inheritance mechanisms (for example, see Jablonka and Lamb, 2005) which interact with each other as well as with environmental and societal factors. For simplicity, in our taxonomy we focus on three broad categories: biological transmission, cultural transmission, and the interaction of biological and cultural transmission (dual transmission)

Biological transmission includes genetic transmission. Individuals inherit nuclear DNA from their parents. Humans also inherit mitochondrial DNA (mtDNA) only from their mothers  Mito-

chondrial DNA codes for the genes of the cell structures which convert food into useable energy, while nuclear DNA codes for the rest of the human genome. The measures of genetic distance used previously in this chapter are based on differences in the distribution of nuclear DNA across populations - that is, on differences in DNA inherited from both parents. As already mentioned, genetic distance is based on neutral genes, which change randomly and are not affected by natural selection. Other parts of the DNA code for genes that are affected by natural selection, such as those affecting eye color or skin color. All these traits are transmitted biologically.

However, genetic transmission is not the only form of biological transmission. In recent years biologists have also given much attention to epigenetic inheritance systems. Epigenetics refers to the mechanisms through which cells with the same genetic information (i.e., DNA) acquire different phenotypes (i.e., observable characteristics) and transmit them to their daughter cells. Examples of epigenetic markers are methylation patterns: DNA methylation is a biochemical process that stably alters the expression of genes in cells by adding a methyl group to a DNA nucleotide. There is currently a debate in the scientific literature about the extent to which epigenetic changes can be inherited from one generation to the next - for instance, see Chandler and Alleman (2008) and Morgan and Whitelaw (2008). An example of possible intergenerational epigenetic inheritance, mentioned by Morgan and Whitelaw (2008), is the Dutch Famine Birth Cohort Study by Lumey (1992), reporting that children born during famine in World War II were smaller than average and that the effects could last two generations (but see also Stein and Lumey, 2002). In principle, epigenetic mechanisms could explain rapid biological changes in populations that could not be due to genetic selection. Epigenetic mechanisms have recently been emphasized by microeconomists working on human capital formation, such as Cunha and Heckman (2007, p. 32), who wrote: "the nature versus nurture distinction is obsolete. The modern literature on epigenetic expression teaches us that the sharp distinction between acquired skills and ability featured in the early human capital literature is not tenable."

Of course, biological inheritance is not the only mode of intergenerational transmission of traits across human beings. Many traits are transmitted culturally from one generation to the next. An important example is the specific language that each child acquires through learning and imitation, usually (but not necessarily) from parents or other close relatives. Other cultural traits include values, habits, and norms. In general, culture is a broad concept, which encompasses a vast range of traits that are not transmitted biologically across generations. The Webster's Encyclopedic

Unabdidged Dictionary defines culture as including "the behaviours and beliefs characteristic of a particular social, ethnic or age group" and "the total ways of living built up by a group of human beings and transmitted from one generation to the other." Richerson and Boyd (2005, p. 5), two leading scholars in the field of cultural evolution, define culture as "information capable of affecting individuals' behavior that they acquire from other members of their species through teaching, imitation, and other forms of social transmission."

Following Jablonka and Lamb (2005), we can distinguish between two forms of cultural transmission, both involving social learning: behavioral transmission and symbolic transmission. Behavioral transmission takes place when individuals learn from each other by direct observation and imitation. Symbolic transmission instead is about learning by means of systems of symbols - for example, by reading books. Most scholars of human evolution believe that the bulk of observed human variation in intergenerationally transmitted traits is mainly due to cultural transmission rather than to biological transmission. For instance, prominent anthropologists Henrich and McElreath (2003, p. 123) write: "While a variety of local genetic adaptations exist within our species, it seems certain that the same basic genetic endowment produces arctic foraging, tropical horticulture, and desert pastoralism [...]. The behavioral adaptations that explain the immense success of our species are cultural in the sense that they are transmitted among individuals by social learning and have accumulated over generations. Understanding how and when such culturally evolved adaptations arise requires understanding of both the evolution of the psychological mechanisms that underlie human social learning and the evolutionary (population) dynamics of cultural systems."

In sum, our classification of modes of intergenerational transmission includes two broad categories: biological transmission (both genetic and epigenetic) and cultural transmission (behavioral and symbolic). However, these two forms of transmission should not be viewed as completely distinct and independent. On the contrary, a growing line of research stresses that human evolution often proceeds from the interaction between biological and cultural inheritance systems, where each system is influenced by the other system. According to Richerson and Boyd (2005, p. 194), genes and culture can be seen as "obligate mutualists, like two species that synergistically combine their specialized capacities to do things that neither can do alone. [...] Genes, by themselves can't readily adapt to rapidly changing environments. Cultural variants, by themselves, can't do anything without brains and bodies. Genes and culture are tightly coupled but subject to evolutionary forces that tug behavior in different directions." This approach to evolution is known as dual in-

heritance theory or gene-culture coevolution (Cavalli-Sforza and Feldman, 1976, 1981; Boyd and Richerson, 1985; Richerson and Boyd, 2005). In such a framework, observable human outcomes can be viewed as stemming from the interplay of genetically and culturally transmitted traits. A well-known example of gene-culture coevolution is the spread of the gene controlling lactose absorption in adults in response to cultural innovations, such as domestication and dairying (Simoons, 1969, 1970; Richerson and Boyd, 2005, chapter 6). The ability to digest milk as an adult (i.e., to be "lactase persistent") is given by a gene that is unequally distributed among different populations: it is prevalent among populations of European descent, but very rare among East Asians and completely absent among Native Americans. It is well-understood that such gene did spread rapidly after the introduction of domestication among populations that kept milk-producing animals, such as cows or goats, reinforcing the advantages from those practices from an evolutionary perspective. In general, dual inheritance - the third "mode of transmission" in our taxonomy - captures such a complex interaction between genetic and cultural factors.

## 5.3  Modes of Operation

Traits can be transmitted from one generation to the next biologically, culturally, or through the interaction of genes and culture (dual transmission). But how do such traits affect economic outcomes? Our taxonomy distinguishes between direct effects and barrier effects.

**Direct Effects.**   Most of the economic literature has focused on *direct* effects of vertically transmitted traits on income and productivity. Such effects occur when individuals inherit traits that directly impact economic performance, either positively or negatively. For example, most contributions on the relation between cultural values and economic development stress inherited norms and beliefs that directly lead to positive or negative economic outcomes. Max Weber (1905), the great German sociologist and political economist, in his classic book *The Protestant Ethic and the Spirit of Capitalism,* provided a systematic and influential study emphasizing the direct positive effects of specific culturally transmitted traits on economic performance. Weber was in part reacting to the Marxist view, which considered cultural beliefs and values, such as religion, as the by-product of underlying economic factors. Instead, Max Weber argued for direct causal effects of culturally transmitted traits on economic outcomes. Specifically, he proposed that the emergence of a new Protestant ethic, which linked "good works" to predestination and salvation, had a direct effect on the rising of the "spirit of capitalism," a new attitude towards the pursuit of economic

prosperity. Among Weber's more recent followers is, for example, the economic historian David Landes (1998, 2000), who titled one of his contributions "Culture Makes Almost All the Difference," and opened it with the line "Max Weber was right." Landes' emphasis was also on the direct economic effects of culture, defined as "the inner values and attitudes that guide a population." According to Landes (p. 12): "This is not to say that Weber's "ideal type" of capitalist could be found only among Calvinists [...]. People of all faiths and no faith can grow up to be rational, diligent, orderly, productive, clean, and humourless. [...] Weber's argument, as I see it, is that in sixteenth- to eighteenth-century northern Europe, religion encouraged the appearance in numbers of a personality type that had been exceptional and adventitious before and that this type created a new economy (a new mode of production) that we know as (industrial) capitalism."

An extensive empirical literature has attempted to directly test Weber's hypotheses, often concluding with a negative assessment of direct effects of Protestant values on economic outcomes. Recent contributors to this literature were Sascha Becker and Ludger Woessman (2009), who used county-level data from nineteenth century Prussia, and attempted to estimate the causal effect of Protestantism on economic performance by exploiting the fact that the Lutheran Reform expanded concentrically from Wittenberg, Martin Luther's city. They concluded that Protestantism fostered economic development, but that the main channel was not the spread of a new work ethic associated with religious values, but the expansion of literacy as a consequence of education in reading the Bible.

The direct effects of religious beliefs on economic outcomes were investigated empirically by Barro and McCleary (2003). Barro and McClearly used instrumental variables, such as the existence of a state religion and of a regulated market structure, to identify the direct effect of religion on growth. They concluded that economic growth is positively associated with the extent of religious beliefs, such as those in hell and heaven, but negatively associated to church attendance. They interpreted their results as consistent with a direct effect of religion - a culturally transmitted set of beliefs - on individual characteristics that foster economic performance. Guiso, Sapienza and Zingales (2003) also studied the effects of religious beliefs on economic attitudes and outcomes, such as cooperation, legal rules, thriftiness, the market economy, and female labor participation. They found that religious beliefs tend to be associated with attitudes conducive to higher income per capita and higher economic growth, and that the effects differ across religious denominations.

While scholars such as Weber have stressed the positive direct effects of cultural traits, such as

the Protestant ethic, other scholars have argued that specific culturally transmitted traits and values can be responsible for economic backwardness and underdevelopment. An influential and widely debated example of this view was provided by the political scientist Edward Banfield (1958) in his classic book *The Moral Basis of a Backward Society*, written in collaboration with his wife Laura Fasano, and based on their visit to the Southern Italian town of Chiaromonte (called "Montegrano" in the book). Banfield argued that the economic backwardness of that society could be partly explained by the direct effects of inherited values summarized by the term "amoral familism," and consisting in a lack of mutual trust and cooperation, and a disregard for the interests of fellow citizens who were not part of one's immediate family. A theory of intergenerational transmission directly inspired by Bansfield's analysis has been provided recently by Guido Tabellini (2008), who also built on Alberto Bisin and Thierry Verdier's (2000, 2001) seminal work on the economics of cultural transmission. In Tabellini's model, parents choose which values to transmit to their children, depending on the patterns of external enforcement and expected future transactions. In particular, Tabellini shows that path dependence is possible: adverse initial conditions can lead to a unique equilibrium where legal enforcement is weak and inherited cultural values discourage cooperation.

A recent example of an empirical study of the direct effects of inherited traits on economic growth is Algan and Cahuc (2010). Algan and Cahuc document how the level of inherited trust of descendants of immigrants in the United States is significantly influenced by the country of origin and the timing of arrival of their ancestors. They then use the inherited trust of descendants of immigrants in the US as a time-varying measure of inherited trust in their country of origin, in order to identify the impact of inherited trust on growth, controlling for country fixed effects. Algan and Cahuc find that changes in inherited trust during the 20th century have a large impact on economic development in a panel of 24 countries.

The above-mentioned contributions are examples of a much larger literature on the direct effects of cultural traits on economic outcomes. There is also a smaller but important literature that has extended the analysis to traits that are transmitted biologically, or stem from the interaction of genes and culture (dual inheritance). An example is the contribution by Galor and Moav (2002), who modeled an intergenerationally transmitted trait affecting humans' fertility strategies. They posited that some individuals inherited traits that induced them to follow a quantity-biased strategy, consisting in the generation of a higher number of children, while other individuals followed

a quality-biased strategy, consisting in the investment of more resources in a smaller number of offspring. Galor and Moav argued that the evolutionary dynamics of these traits had direct implications for the onset of the Industrial Revolution and the following demographic transition. In the preindustrial world, caught in a Malthusian trap, selective pressures favored parental investment, which led to higher productivity. In their model, the spread of this inherited predilection for a smaller number of children led endogenously to the transition out of the Malthusian regime. Galor and Moav in their contribution stressed biological transmission. However, their analysis can also be interpreted as a model of cultural transmission of traits influencing fertility strategies, or as the outcome of the interaction of biological and cultural traits.

A more recent contribution that stresses the direct effects of different distributions of intergenerationally transmitted traits on economic development is Ashraf and Galor (2013a). In that study, Ashraf and Galor focus on genetic diversity. While genetic distance refers to genetic differences between populations, genetic diversity is about heterogeneity within populations. In their study, Ashraf and Galor (2013a) document a non-monotonic relationship between genetic diversity and development, and argue that such relation is causal, stemming from a trade-off between the beneficial and the detrimental effects of diversity of traits on productivity. Again, while the focus of Ashraf and Galor's empirical analysis is on genetic variables, the modes of transmission from intergenerational traits to economic outcomes can operate both through biological and cultural channels, and their interactions. A further discussion of the relation between genetic diversity and ethnic and cultural fragmentation is provided by Ashraf and Galor (2013b).

The interaction of culture and genes is explicitly at the center of the economic analysis of the effecst of lactase persistence provided by Justin Cook (2012). Cook argues that country-level variation in the frequency of lactase persistence is positively and significantly related to economic development in pre-modern times - which he measures by using population density in 1500 CE, as we did earlier in this chapter. Specifically, he finds that an increase in one standard deviation in the frequency of lactase persistent individuals (roughly 24 percentage points) is associated with a 40 percent increase in pre-modern population density. Cook uses instrumental variables (solar radiation) to assess causality, and interprets his results as reflecting the direct effects of inherited cultural and biological traits associated with the introduction of dairying.

**Barrier effects.** As we already mentioned, most of the contributions on the relation between ancestry and economic performance, including the examples mentioned above, tend to focus on

the direct effects of intergenerationally transmitted traits on economic outcomes. However, as we emphasized in the theoretical and empirical analysis presented in the first sections of this chapter, differences in inherited traits can also affect comparative development by acting as barriers to the diffusion of goods, services, ideas and innovations. A focus on barriers can explain why differences in inherited traits may matter, even though many new ideas and innovations are learned "horizontally," from individuals and populations that are not directly related, rather than "vertically," from one's close relatives and ancestors. The fact is that, when barrier effects do exist, vertically transmitted traits also affect horizontal learning and diffusion. People are more likely to learn new ideas and adopt new technologies from other people who, while not directly related to them, share more recent common ancestors and, consequently, also share on average a larger set of inherited traits and characteristics.

The literature on the barrier effects of vertically transmitted traits is not as large as the one on direct effects. In addition to our own contributions, already discussed, a recent example is Guiso, Sapienza and Zingales (2009), who studied the barrier effects of cultural traits by using data on bilateral trust between European countries. They found that bilateral trust is affected by cultural aspects of the match between trusting country and trusted country, such as their history of conflicts and their religious, genetic, and somatic similarities. Lower bilateral trust then acts as a cultural barrier: it is associated with less bilateral trade, less portfolio investment, and less direct investment between the two countries, even after controlling for other characteristics of the two countries. These findings suggest that culturally transmitted traits can have a significant barrier effect on economic interactions between different societies.

Another study that documents the effects of cultural barriers on trade is provided by Gabriel Felbermayr and Farid Toubal (2010). Felbermayr and Toubal measure cultural proximity or distance between countries using bilateral score data from the Eurovision Song Contest, a popular European television show. For instance, viewers in Cyprus award Greek singers more points on average than the Greeks receive from viewers in other countries, and vice versa. In contrast, Cypriot and Turkish viewers give each other below-average scores. Felbermayr and Toubal exploit the variation of these scores within-pair and across time to estimate the effects of cultural proximity on bilateral trade, finding significant effects.

An open question concerns the relationship between direct and barrier effects. Of course, in principle both modes of operation can be at work simultaneously, and some specific traits can play

a role along both channels. For example, populations that inherit values and beliefs that make them more open to risk and innovation could benefit directly from such traits, but may also face lower barriers to interactions with other groups. In general, the study of barrier effects stemming from historical and cultural divergence is a promising area of research, still in its infancy, both from a theoretical and empirical perspective. The taxonomy and discussion presented in this chapter are only a first step towards a more complete understandiing of this important topic.

# 6   Conclusion

In this chapter we provided a theoretical framework and empirical evidence to shed light on a fundamental question: What barriers prevent the diffusion of the most productive technologies from the technological frontier to less developed economies?

In the first part of this chapter, we presented a simple analytical framework to illustrate two basic ideas. The first idea was that genetic distance between populations, which measures their degree of genealogical relatedness, can be interpreted as a summary metric for average differences in traits that are transmitted with variation from one generation to the next. We modeled the transmission of such "vertical" traits - that is, the transmission of characteristics which are passed on vertically across generations within a population over the very long run - and derived the relation between divergence in vertical traits and genetic distance. The second idea was that differences in vertically transmitted traits act as obstacles to horizontal learning and imitation across different populations. We argued that populations that share a more recent common history and are therefore closer in terms of vertical traits tend to face lower costs and barriers to adopting each other's technological innovations.

In the second part of this chapter we brought these ideas to the data. We introduced measures of genetic distance between populations, and used them to test our barrier model of diffusion. We found that, as the model predicts, genetic distance measured relative to the world's technological frontier trumps absolute genetic distance as an explanation for bilateral income differences and for the different usage of specific technological innovations. This was the case both historically, when we measured technological usage on the extensive margin, and for more recent technological developments, when we measured technological usage along the intensive margin. We also documented that, as implied by our model, the effect of genetic distance was more pronounced after a major innovation, such as the onset of the Industrial Revolution, and declined as more populations

adopted the frontier's innovation. Overall, we found considerable evidence that barriers introduced by historical separation between populations have played a key role in the diffusion of technological innovations and economic growth.

In the third and final part of this chapter, we discussed our hypotheses and results within the broader context of the growing literature on the deep historical roots of economic development. To organize our discussion we presented a taxonomy based on Spolaore and Wacziarg (2013). The taxonomy provided a conceptual basis for discussing how intergenerationally transmitted traits could conceivably affect economic outcomes. Our taxonomy distinguished possible economic effects of vertical traits along two dimensions. The first dimension referred to the *mode of transmission* of vertical traits, which could be biological (genetic or epigenetic), cultural (behavioral or symbolic), or resulting from the interaction of genes and culture (dual inheritance). The second dimension defined the *mode of operation* of these traits, depending on whether they have direct effects on economic outcomes, or operate as barriers to economic interactions between populations. We briefly reviewed examples of economic contributions that focused on different effects - direct effects or barrier effects - of traits transmitted biologically, culturally, or through dual transmission. We argued that most of the literature so far has mainly focused on direct effects, while much less attention has been given to the study of barriers to development stemming from long-term cultural and historical divergence.

The topic of human barriers introduced by historical divergence and their effects on social, political and economic outcomes is an exciting emerging field of study. Our own work continues to explore the effects of variation in human relatedness on a variety of political economy outcomes. For instance, Spolaore and Wacziarg (2012b) examines the effects of genealogical relatedness on the propensity for interstate militarized conflict, finding that a smaller genetic distance is associated with a significantly higher probability of a bilateral conflict between two countries. This effect, again, is interpreted as evidence of a barrier between societies characterized by distinct norms, values, preferences and cultures. This time, however, the barrier impedes a costly rather than a beneficial interaction. In ongoing work, we explore the effects of relatedness on trade and financial flows across countries. Finally, we have recently begun an effort to better characterize what genetic relatedness captures, by investigating the relationship between various measures of cultural differences and genetic distance - the goal being to more clearly identify the source of the barriers introduced by a lack of genealogical relatedness. For instance, the barriers could take the form of a lack of trust, differences in preferences or norms, or transactions costs linked to an inability to

communicate and coordinate. This chapter provides only an introduction and first step towards a more comprehensive and systematic analysis of such important, unexplored, and promising topics.

# References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117 (4): 1231-1294.

Aghion, Philippe and Peter Howitt (2008), *The Economics of Growth*, Cambridge, MA: The MIT Press.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg (2003), "Fractionalization," *Journal of Economic Growth*, vol. 8, pp. 55-194.

Algan, Yann and Pierre Cahuc (2010), "Inherited Trust and Growth," *American Economic Review*, 100 (5), 2060-92

Ashraf, Quamrul and Oded Galor (2011), "Dynamics and Stagnation in the Malthusian Epoch." *American Economic Review* 101 (5): 2003-2041.

Ashraf, Quamrul and Oded Galor (2013a), "The 'Out-of-Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development," *American Economic Review*, vol. 103 no. 1, pp. 1-46.

Ashraf, Quamrul and Oded Galor (2013b), "Genetic Diversity and the Origins of Cultural Fragmentation," *American Economic Review,* vol. 103, no. 3, pp. 528-33.

Banfield, Edward C (1958), *The Moral Basis of a Backward Society*, New York, NY: The Free Press

Barro, Robert J. and Rachel McCleary (2003), "Religion and Economic Growth," *American Sociological Review*, 68 (5), 760-81

Barro, Robert J. and Xavier Sala-i-Martin (1997), "Technological Diffusion, Convergence and Growth", *Journal of Economic Growth*, vol. 2, no.1, pp. 1-26.

Barro, Robert J. and Xavier Sala-i-Martin (2003), *Economic Growth* (2nd ed.), Cambridge, MA: MIT Press.

Becker, Sascha O. and Ludger Woessmann (2009), "Was Weber Wrong? A Human Capital Theory of Protestant Economic History," *Quarterly Journal of Economics*, 124 (2), 531-96.

Bisin, Alberto and Thierry Verdier (2000), "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits," *Quarterly Journal of Economics*, 115, 955-88.

Bisin, Alberto and Thierry Verdier (2001), "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory,* 97, 298-319

Bossert, Walter, Conchita D'Ambrosio and Eliana La Ferrara (2011), "A Generalized Index of Fractionalization", *Economica*, 78(312), 723-750.

Boyd, Robert and Peter J. Richerson. 1985. *Culture and the Evolutionary Process*, Chicago: University of Chicago Press.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2006), "Robust Inference with Multi-Way Clustering," *NBER Technical Working Paper* #T0327.

Caselli, Francesco (2005), "Accounting for Cross-Country Income Differences." in *Handbook of Economic Growth*, Vol. 1A, edited by Philippe Aghion and Steven N. Durlauf, pp. 679–741. New York: North-Holland.

Cavalli-Sforza, Luigi Luca and Marcus W. Feldman. 1981. *Cultural Transmission and Evolution: A Quantitative Approach.* Princeton: Princeton University Press.

Cavalli-Sforza, Luigi L., Paolo Menozzi and Alberto Piazza (1994), *The History and Geography of Human Genes*, Princeton: Princeton University Press.

Chanda, Areendam, C. Justin Cook and Louis Putterman (2013), "Persistence of Fortune: Accounting for Population Movements, There Was No Post-Columbian Reversal", *working paper*, Brown University, February.

Chandler V and Alleman M. (2008), "Paramutation: epigenetic instructions passed across generations", *Genetics*, vol. 178, no. 4, pp: 1839-44.

Comin, Diego and Bart Hobijn (2009), "The CHAT Dataset", Harvard Business School Working Paper # 10-035.

Comin, Diego, and Bart Hobijn (2010), "An Exploration of Technology Diffusion," *American Economic Review*, vol. 100, no. 5, December, 2031-59.

Comin, Diego, Bart Hobijn and Emilie Rovito (2008), "World Technology Usage Lags," *Journal of Economic Growth*, vol. 13, no. 4.

Comin, Diego, William Easterly and Erick Gong (2010), "Was the Wealth of Nations Determined in 1000 B.C.?" *American Economic Journal: Macroeconomics*, vol. 2, no. 3, pp. 65-97.

Cook, C. Justin (2012), "The Role of Lactase Persistence in Precolonial Development ", *working paper*, Yale University, August.

Cunha, Flavio and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2): 31-47.

Desmet, Klaus, Michel Le Breton, Ignacio Ortuño-Ortín, and Shlomo Weber (2011), "The Stability and Breakup of Nations: a Quantitative Analysis." Journal of Economic Growth 16: 183-213.

Diamond, Jared. 1997. *Guns, Germs and Steel: The Fate of Human Societies.* New York: Norton & Co.

Fagerberg, Jan (2004), "Innovation: A Guide to the Literature," Chapter 1 in *Oxford Handbook of Innovation*, edited by J. Fagerberg, D. C. Mowery, and R. R. Nelson, Oxford: Oxford University Press.

Galor, Oded and Omer Moav (2002), "Natural Selection and the Origin of Economic Growth," *Quarterly Journal of Economics*, 117 (4), 1133-91.

Glaeser, Edward L., Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer. 2004. "Do Institutions Cause Growth?" *Journal of Economic Growth* 9 (3): 271-303.

Greenberg, J. H. (1956). "The measurement of linguistic diversity." *Language*, 32, 109-15.

Guiso, Luigi, Paola Sapienza and Luigi Zingales (2003), "People's Opium? Religion and Economic Attitudes," *Journal of Monetary Economics* 50 (1), 225-82.

Guiso, Luigi, Paola Sapienza and Luigi Zingales (2009), "Cultural Biases in Economic Exchange," *Quarterly Journal of Economics*, 124 (3), 1095-131.

Gelbermayr, Gabriel J. and Farid Toubal (2010), "Cultural Proximity and Trade," *European Economic Review*, 54 (2), 279-93

Greenberg, Joseph E. (1956), "The Measurement of Linguistic Diversity", *Language*, vol. 32, no.1, pp. 109-115.

Hall, Robert E. and Charles I. Jones (1999), "Why do some countries produce so much more output per worker than others?", *Quarterly Journal of Economics*, vol. 114, no. 11, pp. 83-116

Henrich, Joseph and Richard McElreath (2003), "The Evolution of Cultural Evolution." *Evolutionary Anthropology* 12: 123-135.

Hsieh, Chang-Tai and Peter J. Klenow (2010), "Development Accounting," *American Economic Journal: Macroeconomics*, vol. 2, no. 1, pp. 207–223.

Jablonka, Eva and Marion J. Lamb (2005), *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life.* Cambridge (MA): MIT Press.

Kimura, Motoo (1968), "Evolutionary Rate at the Molecular Level", *Nature*, vol. 217, pp. 624-626.

Landes, David (1998), *The Wealth and Poverty of Nations*, New York: Norton.

Landes, David (2000), "Culture Makes almost All the Difference," in Lawrence E. Harrison and Samuel P. Huntington (eds), *Culture Matters. How Values Shape Human Progress*, New York, NY, USA: Basic Books, 2-13

Lumey, Lambert H. (1992), "Decreased Birthweights in Infants after Maternal in Utero Exposure to the Dutch Famine of 1944-1945." *Paediatric and Perinatal Epidemiology.* 6: 240–253.

Maddison, Angus (2003), *The World Economy: Historical Statistics,* Paris, France: OECD Development Center.

Morgan, Daniel K. and Emma Whitelaw (2008), "The Case for Transgenerational Epigenetic Inheritance in Humans." *Mammalian Genome*, 19: 394-397.

Olsson, Ola and Douglas A. Hibbs, Jr (2005), "Biogeography and Long-Run Economic Development," *European Economic Review,* vol. 49, no. 4, pp. 909-38.

Parente, Stephen L. and Edward C. Prescott (1994), "Barriers to Technology Adoption and Development," *Journal of Political Economy*, vol. 102, no. 2, pp. 298-321.

Parente, Stephen L. and Edward C. Prescott (2002), *Barriers to Riches*, Cambridge: MIT Press

Putterman, Louis and David N. Weil (2010), "Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality." *Quarterly Journal of Economics* 125 (4): 1627-1682.

Richerson, Peter J., and Robert Boyd (2005), *Not by Genes Alone: How Culture Transformed Human Evolution.* Chicago: University of Chicago Press.

Rogers, E. M. (1962, 1995), *Diffusion of Innovations* (4th ed.). New York: Free Press.

Romer, Paul M. (1990), "Endogenous Technological Change", *Journal of Political Economy*, vol. 98, no. 5, pp. S71-S102.

Simoons, Frederick J. (1969), "Primary Adult Lactose Intolerance and the Milking Habit: A Problem in Biological and Cultural Interrelations: I. Review of the Medical Research." *The American Journal of Digestive Diseases* 14: 819-836.

Simoons, Frederick J. (1970), "Primary Adult Lactose Intolerance and the Milking habit: A Problem in Biological and Cultural Interrelations: II. A Culture Historical Hypothesis." *The American Journal of Digestive Diseases* 15: 695-710.

Spolaore Enrico (2014), "Introduction" to *Culture and Economic Growth,* The International Library of Critical Writings in Economics Series, Cheltenham: Edward Elgar, forthcoming.

Spolaore Enrico and Romain Wacziarg (2009), "The Diffusion of Development," *Quarterly Journal of Economics*, vol. 124, no. 2, pp. 469-529.

Spolaore Enrico and Romain Wacziarg (2012a), "Long-Term Barriers to the International Diffusion of Innovations", in Jeffrey Frankel and Christopher Pissarides, eds., *NBER International Seminar on Macroeconomics 2011*, Chapter 1, pp. 11-46. Chicago: University of Chicago Press.

Spolaore Enrico and Romain Wacziarg (2012b), "War and Relatedness", *working paper*, Tufts University and UCLA, June 2012.

Spolaore Enrico and Romain Wacziarg (2013), "How Deep Are the Roots of Economic Development?", *Journal of Economic Literature*, 51(2), June.

Stein, Aryeh D. and Lambert H. Lumey (2002), "The Relationship between Maternal and Offspring Birth Weights after Maternal Prenatal Famine Exposure: the Dutch Famine Birth Cohort Study." *Human Biology* 72: 641-654.

Tabellini, Guido (2008), "The Scope of Cooperation: Values and Incentives," *Quarterly Journal of Economics*, 123 (3), 905-50.

Weber, Max ([1905, 1930] 2005), *The Protestant Ethic and the Spirit of Capitalism*, [translated from the German by Talcott Parsons], London and New York: Routledge.

# Appendix 1 - Technologies used in the various datasets

### A. 24 Technologies in the CEG 1500 AD Dataset.

1. *Military*: Standing army, cavalry, firearms, muskets, field artillery, warfare capable ships, heavy naval guns, ships (+180 guns).

2. *Agriculture*: Hunting and gathering, pastoralism, hand cultivation, plough cultivation.

3. *Transportation* : Ships capable of crossing the Atlantic Ocean, ships capable of crossing the Pacific Ocean, ships capable of reaching the Indian Ocean, wheel, magnetic compass, horse powered vehicles.

4. *Communications*: Movable block printing, woodblock or block printing, books, paper.

5. *Industry*: Steel, iron.

### B. 9 Technologies in the CEG 2000 AD Dataset.

Electricity (in 1990), Internet (in 1996), PCs (in 2002), cell phones (in 2002), telephones (in 1970), cargo and passenger aviation (in 1990), trucks (in 1990), cars (in 1990), tractors (in 1970).

### C. 33 Technologies in the CHAT dataset for 1990-1999.

1. *Agriculture*: Harvest machines, tractors used in agriculture, metric tons of fertilizer consumed, area of irrigated crops, share of cropland area planted with modern varieties (% cropland), metric tons of pesticides.

2. *Transportation*: civil aviation passenger km, lengths of rail line, tons of freight carried on railways, passenger cars in use and commercial vehicles in use.

3. *Medical*: Hospital beds, DPT immunization before age 1, measles immunization before age 1.

4. *Communications*: Cable TV, cell phones, personal computers, access to the Internet, items mailed/received, newspaper circulation, radios, telegrams sent, mainline telephone lines, television sets in use.

5. *Industry and other*: Output of electricity, KwHr, automatic looms, total looms, crude steel production in electric arc furnaces, weight of artificial (cellulosic) fibers used in spindles, weight of synthetic (non cellulosic) fibers used in spindles, weight of all types of fibers used in spindles, visitor beds available in hotels and elsewhere, visitor rooms available in hotels and elsewhere.

**Table 1 – Summary Statistics for the Main Variables of Interest**

Panel A - Mean and Variation

| Variable | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|
| Difference in log income per capita 2005 | 1.3844 | 0.9894 | 0.0000241 | 4.8775 |
| FST genetic distance relative to the English, 1500 | 0.0710 | 0.0555 | 0 | 0.2288 |
| Weighted FST genetic distance relative to the USA | 0.0612 | 0.0475 | 0 | 0.2127 |
| Weighted FST genetic distance between pairs | 0.1124 | 0.0818 | 0 | 0.3364 |
| Geodesic distance (thousands of km) | 7.1349 | 4.1330 | 0.0105 | 19.9512 |

(10,440 observations)

Panel B - Correlations

| | Difference in log income per capita 2005 | FST genetic distance relative to the English, 1500 | Weighted FST gen. dist. relative to the USA | Weighted FST genetic distance between pairs |
|---|---|---|---|---|
| FST genetic distance relative to the English, 1500 | 0.2745* | 1 | | |
| Weighted FST genetic distance relative to the USA | 0.3226* | 0.6105* | 1 | |
| Weighted FST genetic distance between pairs | 0.1950* | 0.2408* | 0.5876* | 1 |
| Geodesic distance (thousands of km) | 0.0126 | 0.0644* | 0.0899* | 0.3317* |

(*: significant at the 5% level. 10,440 observations)

## Table 2 - Income difference regressions
### (Dependent variable: Difference in log per capita income, 2005)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | **OLS with relative GD** | **OLS with simple GD** | **Horserace between simple and relative GD** | **2SLS with 1500 GD** |
| **FST gen. dist. relative to the USA, weighted** | **6.290** **(1.175)***** | | **6.029** **(1.239)***** | **9.720** **(1.974)***** |
| **FST genetic distance** | | **2.164** **(0.596)***** | **0.275** **(0.541)** | |
| Absolute difference in latitudes | 0.232 (0.245) | 0.559 (0.279)** | 0.255 (0.248) | 0.152 (0.300) |
| Absolute difference in longitudes | -0.025 (0.220) | -0.196 (0.240) | -0.007 (0.213) | 0.238 (0.247) |
| Geodesic Distance | -0.012 (0.026) | -0.008 (0.027) | -0.016 (0.025) | -0.042 (0.028) |
| =1 for contiguity | -0.418 (0.060)*** | -0.495 (0.060)*** | -0.414 (0.061)*** | -0.326 (0.069)*** |
| =1 if either country is an island | 0.174 (0.083)** | 0.143 (0.083)* | 0.174 (0.083)** | 0.211 (0.084)*** |
| =1 if either country is landlocked | 0.008 (0.085) | 0.024 (0.090) | 0.005 (0.087) | -0.029 (0.085) |
| =1 if pair shares at least one sea or ocean | -0.001 (0.067) | 0.028 (0.077) | -0.000 (0.067) | -0.024 (0.078) |
| Constant | 1.022 (0.089)*** | 1.143 (0.086)*** | 1.017 (0.090)*** | 0.891 (0.099)*** |
| Standardized Beta (%) | 30.18 | 10.39 | 28.93 | 46.49 |
| R-Squared | 0.11 | 0.07 | 0.11 | 0.09 |

Two-way clustered standard errors in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%.

All regressions are based on 10,440 observations.

**Table 3 - Income difference regressions, regional controls and sample splits**
**(Dependent variable: Difference in log per capital income in 2005, 1870 for column 3)**

| | (1) Continent dummies | (2) Europe 2005 income | (3) Europe with 1870 income | (4) Excluding Europe | (5) Control for Europeans | (6) Excluding SS Africa |
|---|---|---|---|---|---|---|
| **Fst gen. dist. relative to the USA, weighted** | 3.403 (1.284)** | | | 5.183 (1.232)*** | 5.624 (1.143)*** | 4.851 (1.443)*** |
| **Genetic distance, relative to the English** | | 25.920 (11.724)** | 27.054 (6.557)*** | | | |
| **Abs. difference in the shares of people of European descent** | | | | | 0.626 (0.125)*** | |
| Constant | 1.541 (0.315)** | 0.345 (0.201)* | 0.495 (0.154)*** | 1.006 (0.123)*** | 0.864 (0.097)*** | 0.853 (0.071)*** |
| Observations | 10,440 | 253 | 136 | 6,328 | 10,153 | 5,253 |
| Standardized Beta (%) | 16.27 | 31.28 | 43.62 | 24.99 | 27.15 | 17.12 |
| R-Squared | 0.20 | 0.24 | 0.24 | 0.08 | 0.17 | 0.06 |

Two-way clustered standard errors in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%.

In all regressions, controls are included for: Absolute difference in latitudes, absolute difference in longitudes, geodesic distance, dummy for contiguity, dummy if either country is an island, dummy if either country is landlocked, dummy if pair shares at least one sea or ocean.

Column 1 includes continental dummies defined as follows: both in Asia dummy, both in Africa dummy, both in Europe dummy, both in North America dummy, both in Latin America/Caribbean dummy, both in Oceania dummy, dummy if one and only one country is in Asia, dummy if one and only one country is in Africa, dummy if one and only one country is in Europe, dummy if one and only one country is in North America, dummy if one and only one country is in South America.

51

**Table 4: Pairwise correlations between historical measures of economic development**

| | Relative genetic distance to the English, 1500 | Relative genetic distance to the UK, (contemporary) | Abs. difference in population density, 1500 | Abs. difference in log income, 1820 | Abs. difference in log income, 1870 | Abs. difference in log income, 1913 | Abs. difference in log income, 1960 |
|---|---|---|---|---|---|---|---|
| Relative genetic distance to the UK, weighted (contemporary) | 0.6205* (10,585) | 1 (10,585) | | | | | |
| Abs. difference in population density, 1500 | 0.1594* (10,153) | 0.0461* (10,153) | 1 (10,153) | | | | |
| Abs. difference in log income, 1820 | 0.1763* (1,035) | 0.1327* (1,035) | 0.1701* (990) | 1 (1,035) | | | |
| Abs. difference in log income, 1870 | 0.1360* (1,485) | 0.1811* (1,485) | 0.1125* (1,378) | 0.6117* (1,035) | 1 (1,485) | | |
| Abs. difference in log income, 1913 | 0.0840* (1,653) | 0.1839* (1,653) | 0.0739* (1,540) | 0.5411* (1,035) | 0.8996* (1,485) | 1 (1,653) | |
| Abs. difference in log income, 1960 | 0.2347* (4,753) | 0.3229* (4,753) | 0.1242* (4,560) | 0.4018* (820) | 0.6154* (1,035) | 0.7201* (1,176) | 1 (4,753) |
| Abs. difference in log income, 2005 | 0.2745* (10,440) | 0.3228* (10,440) | 0.1173* (10,011) | 0.3297* (990) | 0.4722* (1,431) | 0.4844* (1,596) | 0.6199* (4,753) |

(*: Significant at the 5% level; # of obs. in parentheses)

**Table 5 - Regressions using Historical Income Data**

| | (1) Year 1500 population density | (2) Income 1820 | (3) Income 1870 | (4) Income 1913 | (5) Income 1960 | (6) Income 2005 |
|---|---|---|---|---|---|---|
| Relative Fst genetic distance to the UK, 1500 match | 29.751 (7.168)*** | | | | | |
| Relative Fst genetic distance to the UK, weighted | | 0.671 (0.344)* | 1.691 (0.836)** | 1.984 (0.907)** | 3.472 (0.783)** | 5.075 (0.941)** |
| Constant | 6.693 (0.981)*** | 0.313 (0.063)** | 0.365 (0.076)** | 0.421 (0.064)** | 0.478 (0.077)** | 1.017 (0.088)** |
| Observations | 10,153 | 1,035 | 1,485 | 1,653 | 4,753 | 10,440 |
| Standardized beta (%), maximal sample | 17.77 | 8.75 | 15.02 | 15.02 | 28.82 | 30.58 |
| Standardized beta, restricted sample* | - | 9.89 | 16.30 | 17.36 | 11.15 | 7.49 |
| R-Squared | 0.07 | 0.22 | 0.16 | 0.17 | 0.17 | 0.11 |

Two-way clustered standard errors in parentheses; * significant at 10%; ** significant at 5%; ** significant at 1%

In all regressions, controls are included for: Absolute difference in latitudes, absolute difference in longitudes, geodesic distance, dummy for contiguity, dummy if either country is an island, dummy if either country is landlocked, dummy if pair shares at least one sea or ocean.

Population density data for 1500 are from McEvedy and Jones (1978). Income data for 1820, 1870 and 1913 are from Maddison (1994). Income data for 1960 and 2005 are from the Penn World Tables.

* The restricted sample for columns (2)–(6) consists of 820 country pairs constructed from 41 countries (Algeria, Australia, Austria, Belgium, Brazil, Canada, China, Denmark, Egypt, Finland, France, Greece, India, Indonesia, Iran, Ireland, Italy, Jamaica, Japan, Jordan, Korea, Malaysia, Mexico, Morocco, Nepal, Netherlands, New Zealand, Norway, Philippines, Portugal, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Syria, Taiwan, Thailand, Turkey, U.S.A, United Kingdom).

**Table 6 - Summary Statistics for genetic distance and technological adoption levels (from the Comin, Easterly and Gong 2010 data)**

Panel A - Mean and variation

| Variable | # of Obs. | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|---|
| Avg. tech. adoption in agriculture in 1970-2000 | 6,105 | 0.1998 | 0.2327 | 0 | 0.8553 |
| Avg. tech. adoption in communications in 1970-2000 | 7,381 | 0.2601 | 0.1894 | 0 | 0.7911 |
| Avg. tech. adoption in transportation in 1970-2000 | 6,441 | 0.1986 | 0.1600 | 0 | 0.8443 |
| Avg. tech. adoption in industry in 1970-2000 | 5,565 | 0.3005 | 0.2153 | 0 | 1.0278 |
| Avg. of the sectoral tech. adoption indexes in 1970-2000 | 7,503 | 0.2129 | 0.1807 | 0 | 0.8378 |
| Absolute difference in log income, 2005 | 10,440 | 1.3844 | 0.9894 | 0 | 4.8775 |
| FST gen. dist. relative to the USA, weighted | 10,585 | 0.0611 | 0.0473 | 0 | 0.2127 |
| Simple FST genetic distance | 10,585 | 0.1126 | 0.0816 | 0 | 0.3364 |

54

Panel B – Pairwise Correlations

| | Avg. tech. adoption in agriculture in 1970-2000 | Avg. tech. adoption in communications in 1970-2000 | Avg. tech. adoption in transportation in 1970-2000 | Avg. tech. adoption in industry in 1970-2000 | Avg. of the sectoral tech. adoption indexes in 1970-2000 | Absolute difference in log income, 2005 | FST gen. distance relative to the USA, weighted |
|---|---|---|---|---|---|---|---|
| Avg. tech. adoption in communications in 1970-2000 | 0.5550* (5,886) | 1 (7,381) | | | | | |
| Avg. tech. adoption in transportation in 1970-2000 | 0.5308* (5,356) | 0.4331* (6,216) | 1 (6,441) | | | | |
| Avg. tech. adoption in industry in 1970-2000 | 0.5396* (5,460) | 0.6335* (5,460) | 0.5192* (5,050) | 1 (5,565) | | | |
| Avg. of the sectoral tech. adoption indexes in 1970-2000 | 0.7615* (5,995) | 0.7010* (7,260) | 0.7591* (6,441) | 0.7735* (5,565) | 1 (7,503) | | |
| Absolute difference in log income, 2005 | 0.4106* (6,105) | 0.5619* (7,381) | 0.4662* (6,441) | 0.7210* (5,565) | 0.6521* (7,503) | 1 (10,440) | |
| FST gen. dist. relative to the USA, weighted | 0.1301* (6,105) | 0.1877* (7,381) | 0.1248* (6,441) | 0.2958* (5,565) | 0.1975* (7,503) | 0.3226* (10,440) | 1 (10,585) |
| Simple FST genetic distance | -0.0562* (6,105) | 0.0862* (7,381) | -0.0409* (6,441) | 0.1407* (5,565) | 0.0042 (7,503) | 0.1950* (10,440) | 0.5859* (10,585) |

*: Significant at the 5% level; # of obs. in parentheses.

**Table 7 - Technological Distance and Genetic Distance in the contemporary period (1970-2000)**
**(Dependent variables: Measures of technological usage from Comin, Easterly and Gong, as described in row 2)**

| | (1) Agricultural Technology | (2) Communications Technology | (3) Transportation Technology | (4) Industrial Technology | (5) Overall Technology |
|---|---|---|---|---|---|
| **FST gen. dist. relative to the USA, weighted** | **0.402** (0.268) | **0.500** (0.212)** | **0.608** (0.185)*** | **1.149** (0.288)*** | **0.745** (0.216)*** |
| Absolute difference in latitudes | 0.687 (0.121)*** | 0.274 (0.066)*** | 0.306 (0.057)*** | 0.329 (0.081)*** | 0.361 (0.082)*** |
| Absolute difference in longitudes | 0.405 (0.129)*** | 0.089 (0.055) | 0.305 (0.072)*** | 0.174 (0.069)** | 0.243 (0.088)*** |
| Geodesic Distance | -0.050 (0.014)*** | -0.016 (0.006)** | -0.036 (0.008)*** | -0.024 (0.007)*** | -0.032 (0.010)*** |
| =1 for contiguity | -0.050 (0.014)*** | -0.077 (0.012)*** | -0.053 (0.013)*** | -0.090 (0.018)*** | -0.071 (0.012)*** |
| =1 if either country is an island | 0.118 (0.077) | 0.057 (0.027)** | 0.093 (0.047)** | 0.062 (0.023)*** | 0.116 (0.048)** |
| =1 if either country is landlocked | -0.007 (0.028) | 0.018 (0.017) | -0.008 (0.011) | 0.013 (0.023) | -0.016 (0.014) |
| =1 if pair shares at least one sea or ocean | 0.036 (0.027) | -0.010 (0.015) | 0.014 (0.015) | 0.001 (0.020) | 0.009 (0.019) |
| Constant | 0.089 (0.029)*** | 0.199 (0.018)*** | 0.148 (0.018)*** | 0.198 (0.023)*** | 0.147 (0.018)*** |
| Observations | 6,105 | 7,381 | 6,441 | 5,565 | 7,503 |
| Standardized Beta (%) | 8.38 | 12.73 | 18.68 | 25.97 | 19.81 |
| R-Squared | 0.25 | 0.10 | 0.14 | 0.16 | 0.17 |

Two-way clustered standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

**Table 8 - Technological Distance and Genetic Distance in the Year 1500**

**(Dependent variables: Measures of technological usage from Comin, Easterly and Gong, as described in row 2)**

| | (1) Agricultural Technology | (2) Military Technology | (3) Communications Technology | (4) Transportation Technology | (5) Industrial Technology | (6) Overall Technology |
|---|---|---|---|---|---|---|
| **Relative Fst genetic distance to the English, 1500 match** | **0.551** (0.281)* | **1.752** (0.326)*** | **1.279** (0.288)*** | **1.926** (0.299)*** | **1.673** (0.271)*** | **1.524** (0.229)*** |
| Absolute difference in latitudes | 0.189 (0.096)** | 0.383 (0.094)*** | 0.758 (0.092)*** | 0.172 (0.064)*** | 0.138 (0.061)** | 0.377 (0.065)*** |
| Absolute difference in longitudes | -0.329 (0.082)*** | -0.018 (0.066) | -0.017 (0.068) | -0.039 (0.048) | 0.061 (0.091) | -0.066 (0.061) |
| Geodesic Distance | 0.049 (0.010)*** | 0.009 (0.010) | 0.009 (0.008) | 0.014 (0.007)** | 0.048 (0.010)*** | 0.025 (0.007)*** |
| =1 for contiguity | 0.037 (0.026) | -0.025 (0.019) | -0.042 (0.024)* | -0.006 (0.021) | 0.023 (0.025) | 0.014 (0.014) |
| =1 if either country is an island | -0.049 (0.058) | -0.087 (0.029)*** | -0.095 (0.053)* | -0.073 (0.020)*** | -0.180 (0.031)*** | -0.092 (0.024)*** |
| =1 if either country is landlocked | 0.017 (0.026) | -0.051 (0.018)*** | -0.020 (0.016) | -0.048 (0.011)*** | 0.006 (0.023) | -0.022 (0.011)** |
| =1 if pair shares at least one sea or ocean | -0.006 (0.020) | -0.105 (0.034)*** | -0.018 (0.033) | -0.046 (0.029) | 0.050 (0.029)* | -0.019 (0.025) |
| Constant | 0.082 (0.034)** | 0.166 (0.036)*** | 0.086 (0.026)*** | 0.069 (0.024)*** | -0.126 (0.039)*** | 0.016 (0.020) |
| Observations | 5,253 | 5,886 | 5,886 | 5,253 | 5,253 | 5,886 |
| Standardized Beta (%) | 10.41 | 29.26 | 19.95 | 41.81 | 25.27 | 31.63 |
| R-Squared | 0.23 | 0.27 | 0.36 | 0.32 | 0.46 | 0.44 |

Two-way clustered standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

57

**Table 9 - Bilateral regressions of technological distance on relative genetic distance for 33 technologies (CHAT dataset averaged over 1990-1999)**

| | | Fst gen. dist. relative to the USA, weighted | Observations (countries) | Standardized Beta (%) | R-Squared |
|---|---|---|---|---|---|
| | **Agricultural technologies** | | | | |
| (1) | Harvest machines | 2.044 (1.134)* | 3,486 (84) | 5.91 | 0.17 |
| (2) | Tractors used in agriculture, | 19.615 (8.245)** | 5,778 (108) | 9.05 | 0.25 |
| (3) | Metric tons of fertilizer consumed | 73.393 (23.062)*** | 5,778 (108) | 11.68 | 0.23 |
| (4) | Area of irrigated crops | 0.453 (0.276)* | 5,565 (106) | 7.21 | 0.03 |
| (5) | Share of cropland area planted with modern varieties (% cropland) | 0.182 (0.080)** | 3,321 (82) | 7.20 | 0.02 |
| (6) | Metric tons of pesticides | 0.738 (0.893) | 4,465 (95) | 2.62 | 0.12 |
| | **Transportation technologies** | | | | |
| (7) | Civil aviation passenger km | 0.484 (0.254)* | 3,828 (88) | 11.29 | 0.21 |
| (8) | Lengths of rail line | 0.397 (0.275) | 4,656 (97) | 5.26 | 0.28 |
| (9) | Tons of freight carried on railways | 2.330 (1.421) | 4,005 (90) | 10.63 | 0.16 |
| (10) | Passenger cars in use | 0.245 (0.082)*** | 5,886 (109) | 15.88 | 0.26 |
| (11) | Commercial vehicles in use | 0.066 (0.025)*** | 5,050 (101) | 23.50 | 0.29 |

| | | Fst gen. dist. relative to the USA, weighted | Observations (countries) | Standardized Beta (%) | R-Squared |
|---|---|---|---|---|---|
| | **Medical technologies** | | | | |
| **(12)** | **Hospital beds** | 1.481 | 5,565 | 1.31 | 0.17 |
| | | (4.319) | (106) | | |
| **(13)** | **DPT immunization before age 1** | 0.137 | 5,778 | 3.54 | 0.01 |
| | | (0.156) | (108) | | |
| **(14)** | **Measles immunization before age 1** | 0.141 | 5,778 | 3.71 | 0.01 |
| | | (0.162) | (108) | | |
| | **Communications technologies** | | | | |
| **(15)** | **Cable TV** | 74.485 | 4,753 | 4.23 | 0.16 |
| | | (56.305) | (98) | | |
| **(16)** | **Cell phones** | 0.109 | 5,778 | 8.21 | 0.12 |
| | | (0.044)** | (108) | | |
| **(17)** | **Personal computers** | 0.247 | 4,950 | 12.53 | 0.21 |
| | | (0.099)** | (100) | | |
| **(18)** | **Access to the Internet** | 0.192 | 5,778 | 14.25 | 0.28 |
| | | (0.072)*** | (108) | | |
| **(19)** | **Items mailed/received** | 0.097 | 2,346 | 11.00 | 0.21 |
| | | (0.074) | (69) | | |
| **(20)** | **Newspaper circulation** | 0.245 | 5,886 | 10.43 | 0.25 |
| | | (0.101)** | (109) | | |
| **(21)** | **Radios** | 0.064 | 5,886 | 1.87 | 0.12 |
| | | (0.139) | (109) | | |
| **(22)** | **Telegrams sent** | 0.312 | 2,211 | 5.74 | 0.07 |
| | | (0.260) | (67) | | |
| **(23)** | **Mainline telephone lines** | 0.185 | 5,886 | 11.54 | 0.28 |
| | | (0.067)*** | (109) | | |
| **(24)** | **Television sets in use** | 0.492 | 5,886 | 18.78 | 0.31 |
| | | (0.141)*** | (109) | | |

| | | Fst gen. dist. relative to the USA, weighted | Observations (countries) | Standardized Beta (%) | R-Squared |
|---|---|---|---|---|---|
| **Industrial technologies and other** | | | | | |
| **(25)** | **Output of electricity, KwHr** | 34.477 (13.849)** | 5,565 (106) | 8.16 | 0.23 |
| **(26)** | **Automatic looms** | 0.828 (0.304)*** | 3,570 (85) | 11.19 | 0.06 |
| **(27)** | **Total looms** | 1.200 (0.361)*** | 3,570 (85) | 8.95 | 0.08 |
| **(28)** | **Crude steel production in electric arc furnaces** | 0.091 (0.031)*** | 2,278 (68) | 8.10 | 0.08 |
| **(29)** | **Weight of artificial (cellulosic) fibers used in spindles** | 0.425 (0.354) | 2,145 (66) | 3.89 | 0.10 |
| **(30)** | **Weight of synthetic (non cellulosic) fibers used in spindles** | 2.045 (0.819)** | 2,145 (66) | 9.89 | 0.20 |
| **(31)** | **Weight of all types of fibers used in spindles** | 7.832 (2.759)*** | 2,850 (76) | 12.10 | 0.07 |
| **(32)** | **Visitor beds available in hotels and elsewhere** | 24.245 (7.518)*** | 5,565 (106) | 9.31 | 0.10 |
| **(33)** | **Visitor rooms available in hotels and elsewhere** | 13.518 (3.884)*** | 5,778 (108) | 10.50 | 0.10 |

Two-way clustered standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1% Unless specified in parentheses, the dependent variable is the absolute difference in *per capita* prevalence of the technology between country *i* and country *j*.

All regressions include controls for absolute difference in latitudes, absolute difference in longitudes, geodesic distance, dummy=1 for contiguity, dummy=1 if either country is an island, dummy=1 if either country is landlocked, dummy =1 if pair shares at least one sea or ocean.
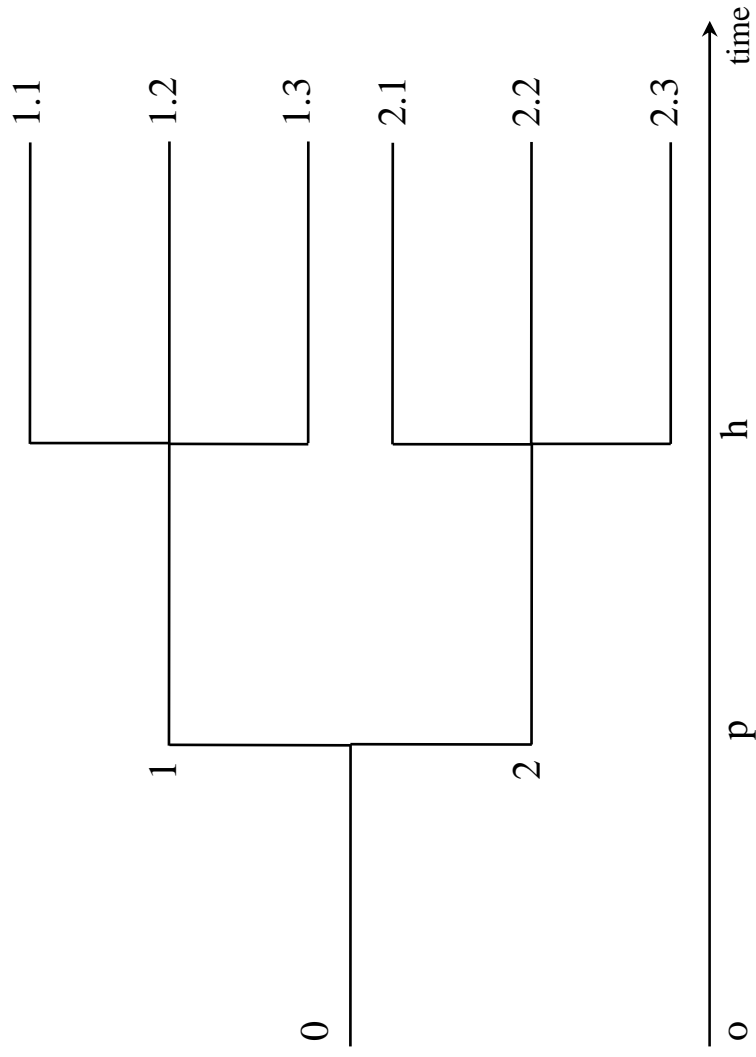
**Figure 1 - Population Tree**

**Figure 2 - Genetic distance among 42 populations.**
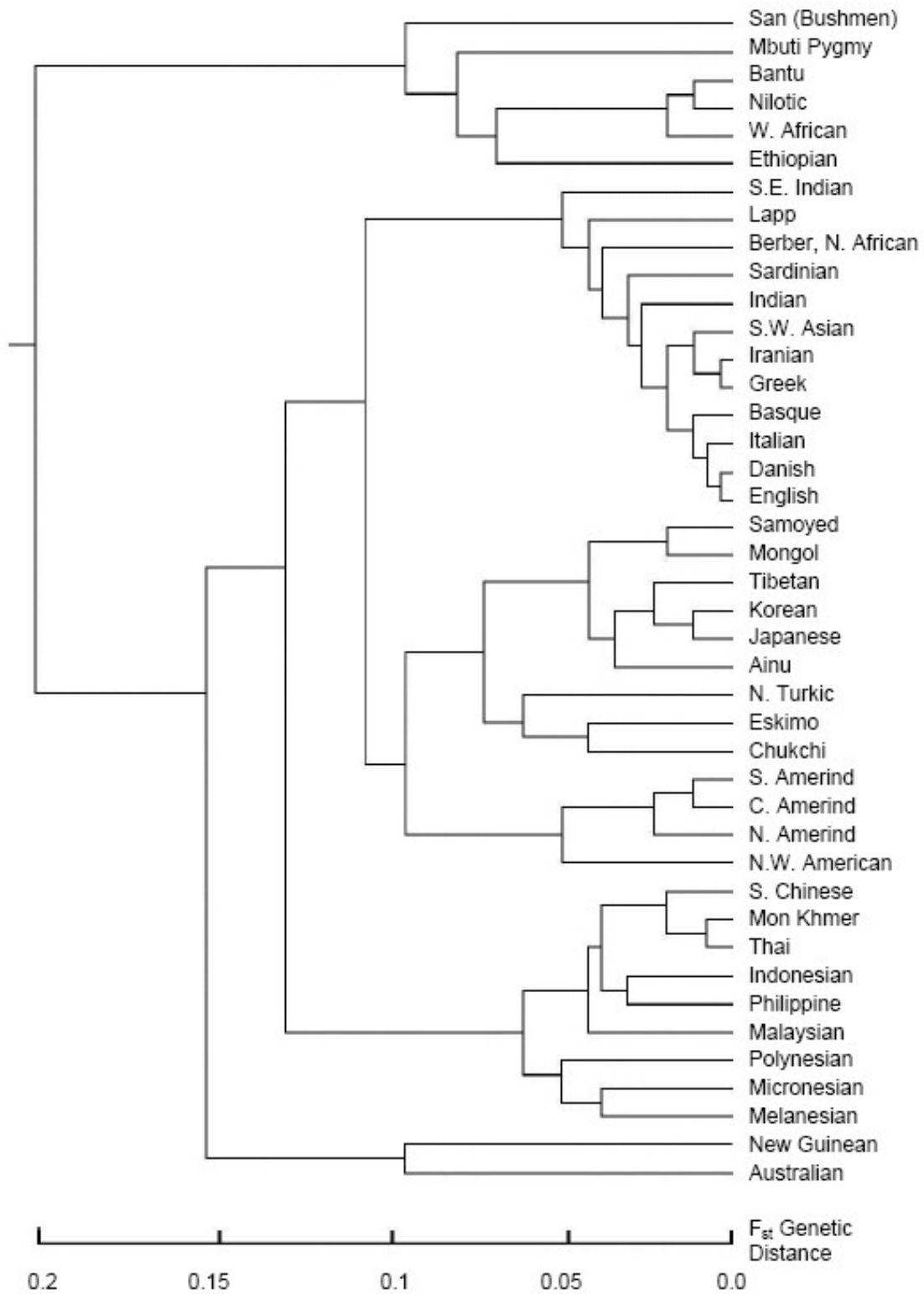**Source: Cavalli-Sforza et al., 1994.**

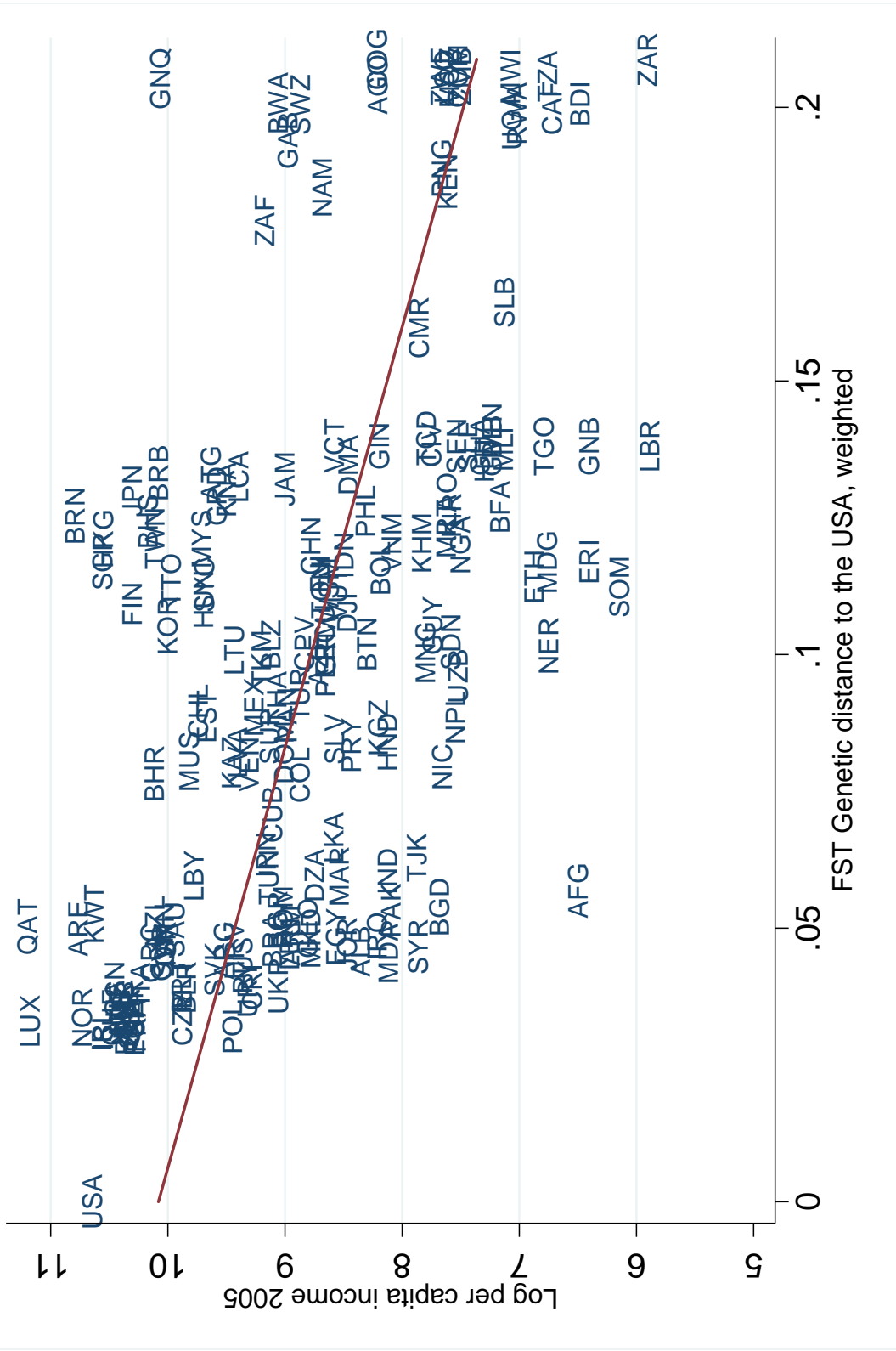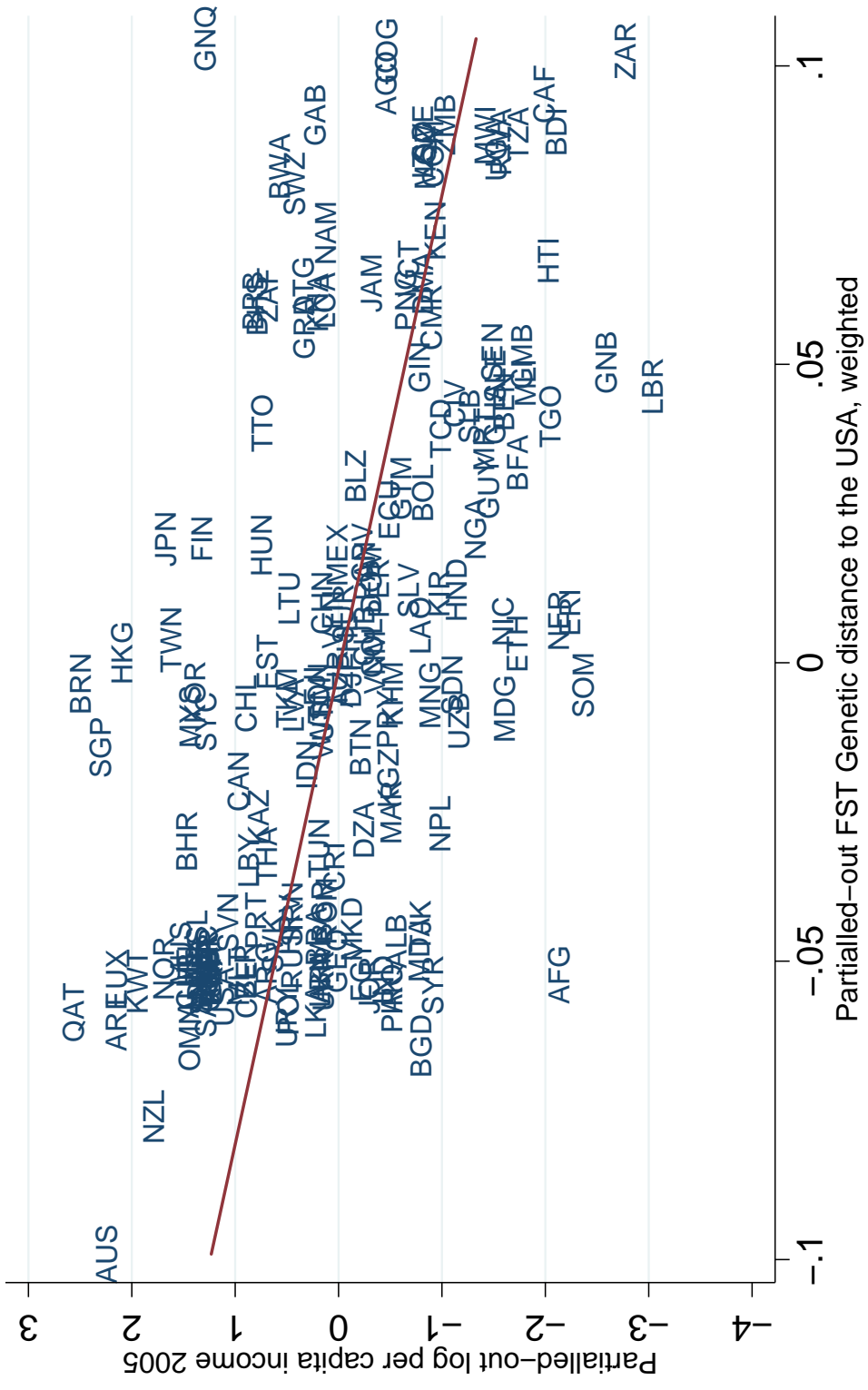Figure 3 – Log Income in 2005 and Genetic Distance to the USA

Figure 4 – Log Income in 2005 and Genetic Distance to the USA, partialling out geodesic distance to the USA

Figure 5 - Standardized effect of genetic distance over time, 1820-2005