A reassessment of the Great Divergence debate: towards a reconciliation of apparently distinct determinants

VICTOR COURT
Science Policy Research Unit (SPRU), Business School, University of Sussex, Jubilee Building, Brighton, BN1 9SL, UK and Chair Energy & Prosperity, Institut Louis Bachelier, 28 place de la Bourse, 75002 Paris, France

This article looks at the most recent data to define when the Little and Great Divergence occurred. It sorts the deep determinants of economic development into three categories (biogeography, culture-institutions, and contingency-conjuncture) to provide a comprehensive review of these factors in the context of the Great Divergence, and it discusses the concepts of persistence and reversal of fortune. The paper concludes that the Great Divergence was never an inevitability but became an increasingly likely prospect as time progressed. Furthermore, biogeography, culture-institutions, and contingency-conjuncture are not contradictory hypotheses. Rather, there is a clear pattern of change over time of the relative importance of these three categories of determinants. Further research is needed to uncover the underlying causal link or latent variable that could explain the successive relative importance over time of biographical, cultural–institutional, and contingent–conjunctural determinants of the Great Divergence.

1. Introduction

1.1. The occurrence and persistence of the Great Divergence

The differential timings of the economic take-off among regions of the world and associated demographic changes have led to the phenomenon called the Great Divergence, which reflects in global differences in economic production levels, generally measured in terms of gross domestic product (GDP) per capita (Figure 1). The link between the industrialization of nations over time and the occurrence of the Great Divergence is one thing, but it is quite another to see the persistence of this phenomenon for decades or even centuries. As a matter of fact, the gap in GDP per capita between the richest and the poorest world regions (Western Offshoots and Africa, respectively) has widened considerably from a modest ratio of 3:1 in 1800 to 15:1 in 2000. This ratio then declined to 10:1 in 2016, showing that there has been some convergence. The examination of national data instead of world regions reveals that global inequality has increased steadily over the past 200 years, appearing to peak around

---

1 In his review of Pomeranz’s (2000) book, Frank (2001) attributes the term “Great Divergence” to Huntington (1996). Strangely enough, there is no trace of the words Great Divergence in Huntington’s (1996) work, so the origin of the term remains a mystery. One possibility is that Kenneth Pomeranz himself came up with the term, perhaps after reading Pritchett’s (1997) article.
Figure 1. The Great Divergence in GDP per capita of world regions, 1800–2016: (a) absolute scale and (b) log scale. Data source: Bolt et al. (2013 2018).

2000, then decreasing up to 2008 before increasing again slightly until 2011 where the estimation stops (Milanovic, 2016, p. 166). To explain the Great Divergence, economic historians used to study three kinds of deep determinants: biogeography, culture, and institutions. More recently, various studies have examined the importance of reactions to local contingent pressures within a global conjuncture to explain the occurrence of the Great Divergence. This work has been salutary in revoking the Eurocentrist bias of many past theories that sought only to find what other parts of the world lacked to explain why they had not industrialised as quickly as Western Europe. Moreover, if economic historians used to rely primarily on narrative analyses (supported by quantitative evidence though), economists have applied their quantitative tools to very long-term analyses during the past two decades. This new and rapidly growing literature has resulted in the design of both theoretical models and econometric assessments of long-term economic growth. Hence, the Great Divergence literature is now composed of thousands of articles and hundreds of books, mixing purely qualitative assessments, theoretical models, and econometric study with increasingly sophisticated designs. However, in spite of these efforts, the phenomenon of Great Divergence seems to remain the deepest mystery of economic history. One could even add that the many hypotheses and analyses for explaining the Great Divergence generate challenges for retaining an accurate “big picture” of the problem.

1.2. Missing perspective, goal, and organization of the paper

Nunn (2014) is a compelling effort to synthesise the literature examining the long-term impacts of historical events, but it does not use the Great Divergence as its key focus.

2 World regions and national averages respectively measure inter-regional and international inequality of the global wealth. Taking into account domestic inequalities allows a more accurate measure of global inequality. In this case, the inequality of wealth between the world’s inhabitants increased sharply from the early nineteenth century to the 1970–1980s. Since then, the global inequality of wealth among all the world’s inhabitants has not shown a clear trend (Milanovic, 2016, p. 120). In summary, there are clear signs of convergence in the international distribution of average per capita wealth across world’s regions, but the convergence is less evident if we consider the average wealth per capita of each nation or the distribution of wealth among all individuals in isolation.

3 I deliberately use the term deep to characterise determinants that are sometimes defined as ultimate in other studies. The reason for this choice is that no single factor, except perhaps the propensity of individuals to act through their daily choices, has the potential to ultimately determine the economic trajectory of a group of people and their descendants during several centuries or millennia.
As a result, Nunn’s (2014) literature review includes elements of no direct relevance to the Great Divergence debate (e.g., the heterogeneous regional impact of the slave trade on the expansion of polygamy), and it omits several references of significant value. More recently, Vries (2016) produced a comprehensive overview of different narrative analyses of the Great Divergence but it disregards recent empirical (econometric) evidence. As a result, Vries’s (2016) review summarizes the current state of controversies within the Great Divergence debate, which is highly valuable, but it does not distinguish between the most important determinants and the more minor factors. Therefore, it fails to provide a clearer understanding of the phenomenon. Hence, there is a need for a comprehensive literature review of all deep determinants of economic growth in the context of the Great Divergence that would fit within one article.

The present article aims to fill this gap and differentiate between the most and the least important determinants of the Great Divergence. For the sake of consistency and space, theoretical models of economic growth in the long run are not investigated in this paper (on this subject, see the unified growth theory of Galor, 2011). Similarly, the role of proximate determinants of modern economic growth for the persistence of the Great Divergence (namely, the role of human capital and institutional barriers to cope with the world technology frontier through technology diffusion and adoption; the mismatch between the technical needs of less economically developed countries and the world technology frontier; and the impact of international trade in raw materials, manufactured goods, and financial assets) are left out of scope (see Acemoglu, 2009 and Aghion & Durlauf, 2005, 2014). However, the recent econometric literature on the occurrence of the Great Divergence will be reviewed along with older qualitative analyses.

Synthesising the knowledge on the occurrence and persistence of the Great Divergence in a single document will help to frame future discussions on this subject and even advance a new agenda for a grand synthesis of the Great Divergence phenomenon. Understanding the Great Divergence requires to account for the co-dynamics of different deep determinants of long-term development, which is already done in several studies (e.g., Bogart, 2011; Kopsidis & Wolf, 2012). But the current paper also shows that the relative importance of different categories of deep determinants changes over time. Accordingly, research should also concentrate on the elaboration of a unified framework able to account for this evolution in order to finally formulate an accurate synthetic explanation of the Great Divergence.

Section 2 frames the Great Divergence phenomenon in broader global history to highlight its uniqueness and timing. The different deep determinants of economic development are analysed in Section 3. Then, Section 4 discusses the notions of persistence and reversal of fortune. Finally, a summary of the contributions of this article is given in Section 5, along with recommendations for future research.

4 Nevertheless, I acknowledge that the literature reviews of Nunn (2014) and Vries (2016) have been very helpful both in adding additional references and in improving the conceptual content of an earlier version of the present article.

5 Empirical investigations of long-run causality have improved, particularly thanks to the use of “Instrumental Variables (IV)” and “two-stages least square (2SL)” methods instead of the more simple “Ordinary Least Square” approaches. For the sake of space and clarity, the technical details of these methods will not be investigated for each empirical study cited in the present article. The reader is invited to consult recent textbooks such as Greene (2008, pp. 314–353) and Wooldridge (2013, pp. 490–528).
2. Past efflorescences and the timings of the Little and Great Divergence

Without further scrutiny, the left part of Figure 1a could lead to two misconceptions, namely that 1. prior to 1800, all societies were stuck in a Malthusian trap with no growth at all, and that consequently 2. all pre-modern societies really looked the same with more developed regions having no distinguishing characteristics. Looking at the data on a log scale in Figure 1b instead of an absolute scale is a first step to demystify these two ideas, and this section details why such thoughts are false by characterising different growth regimes prior to, during, and after the Great Divergence. This is necessary to define the location and timing of the Little and Great Divergence for which different determinants will then be analysed in Sections 3 and 4.

2.1. The importance of past efflorescences for the Great Divergence debate

Prior to the Industrial Revolution, the level of economic and political organisation of societies increased gradually, but neither steadily nor equally. Indeed, technical and political leadership never remained in the same place but changed over time as kingdoms and empires rose and fell in different continents. Furthermore, it makes no doubt that technology improved considerably between the Stone Age and the dawn of the Industrial Revolution. Yet, during all pre-modern times, any increase in food and manufactured goods production generated by technical change or land expansion was primarily channelled toward an increase in the size of the population, providing only a tenuous increase in living conditions relative to modern standards. Hence, intensifying modes of food production—from hunting–gathering to pastoralism, shifting farming, traditional farming, and modern farming—provided support for increasing population densities and sizes (Smil, 2017).

Goldstone (2002) convincingly argues that, prior to the Industrial Revolution, many regions displayed patterns of temporary Smithian efflorescences associated with a marked increase in per-capita incomes, but that growth always petered out eventually and never led towards sustained economic growth. Economic expansions of the High Middle Ages (eleventh to thirteenth centuries) followed by the Renaissance (fourteenth to seventeenth centuries) in Northwestern Europe have received significant attention. As early as the eleventh century, Northern European agriculture was improved by the adoption of wind and water power, heavy plough, and rigid horse-collar. Then, from at least the twelfth century onward, the urban trade centres of Venice, Genoa, the Hanse, and the Rhenish imperial cities blossomed into a trans-European set of urban networks that combined management of urban and rural market-oriented craft production (“proto-industrialisation”) with a flourishing intra-European and extra-European global trade (Goldstone, 2002, p. 329). Song China (c. 960–1280) was probably even more prosperous and technically advanced than the wealthiest parts of Europe in the twelfth century because of its extensive development of water power, iron works, and shipping (Goldstone, 2002, p. 339; 2009, p. 49). Recent empirical data confirm that this example of past efflorescence then petered out (Broadberry et al., 2018).

---

6 See Morris (2010, pp. 183–184) for the cases of Western and Eastern Eurasia’s cores during the past 10,000 years, and Kindleberger (1996) for a more precise analysis of the past 500 years.

7 There are other well-studied past efflorescences such as Classical Greece (fifth and fourth centuries BCE, see Ober, 2015), the Dutch Republic Golden Age (c. 1570–1670, see de Vries, 2000), Tokugawa Japan (c. 1600–1870, see Saito, 2015), and recurring periods of growth spurts followed by bursts in different parts of Africa (from 1690 to 2010, see Jerven, 2010).
To varying degrees, all growth regimes before industrialisation can commonly be characterised as organic-Smithian, i.e., limited and non-sustained growth supported by an increasing division and specialisation of labor and extension of the market fuelled by renewable organic resources derived from solar energy. On the contrary, modern economic growth can be defined as mineral-Schumpeterian, i.e., substantial and sustained growth through a massive use of new fossil energy sources and raw materials, new technologies driven by continuing innovation, and new institutions. An important point is that improved long-run economic performances associated with the progressive transition from organic-Smithian to mineral-Schumpeterian regimes occurs not so much through an increase in the rate of growing, but through a decline in the rate and frequency of shrinking episodes (Broadberry & Wallis, 2017). Therefore, the efflorescence concept and the evidence, in a way, turn the question of the origin of the Great Divergence upside down: the real puzzle is not to find the cause of the early Western take-off lead by England, but rather to understand how England managed to avoid a probable upcoming decline around 1750–1800, and instead created self-sustaining growth after 1830 (Goldstone, 2002, p. 360). Sections 3 and 4 will show that there is indeed no automatic or easy transition from early modern Smithian-limited to fully modern Schumpeterian-sustained economic growth.

2.2. The location and timing of the Little and Great Divergence

As already said, Figure 1b tempers the idea suggested by Figure 1a that the global gap between rich and poor appeared only at the beginning of the nineteenth century and then widened quickly during the twentieth century. Such a view is representative of the California School of Economic History, a term coined by Goldstone (2009, p. viii) to designate scholars from different universities in California who share a common view of the Great Divergence. Members of this paradigm do not see the Great Divergence as the culmination of a long process of dynamic West versus stagnant East, but rather assert that China and Europe shared remarkable similarities of pre-industrial economic expansion based on Smithian dynamics including highly productive agriculture, sophisticated manufacturing, significant cities, and expanded commercial networks (Wong, 1997, p. 278). In the same way, Pomeranz (2000) suggests that circa 1750, the economic cores of Britain, Eastern China (i.e., the Yangtze River delta), and Japan had much more similarities than differences regarding capital accumulation, economic institutions (such as labor markets and the security of property rights), scale and nature of luxury demand, and even material standards of living. Other scholars such as Flynn & Giraldez (1995), Frank (1998), Hobson (2004), Perdue (2005), Goody (2010), and Marks (2015) are more radical and present arguments for the relative backwardness of Europe before industrialisation, and the primordial role that China played in the world economy to enable the Western European take-off. The main argument of this radical wing of the California School rests on the reality of the huge surplus of silver (whose magnitude is still debated) that flowed from the European colonies of Latin America to China in exchange for silk, ceramics, gold, copper cash, and tea exports towards Western Europe. The radical wing of the California School sees in this phenomenon a clear demonstration of the economic hegemony of China, whereas for more nuanced scholars it rather corresponds to the monetisation of the Chinese economy and a financial windfall for Western Europe.
The anti-Eurocentrist work of the California School has been instrumental in reviving the Great Divergence debate, but perhaps at the expense of a tendency to exaggerate the similarities of Western Europe and Eastern Asia (Vries, 2010). Van Zanden (2008) shows that Western Europe did not perform better than Japan and Eastern China in terms of “horizontal” institutions organising citizen interactions. Indeed, price convergence and volatility show that integration levels of labor and grains markets were rather similar in Western Europe and Eastern Asia. However, evidence on interest rates and skill premiums suggest that, as early as the fourteenth century, “vertical” institutions regulating the relationships between the state and citizens (i.e., capital markets and property rights), performed far better in Western Europe compared to Eastern China and Japan. Consistent with this idea, if the California School still supports that “the West and the Rest” (to take Huntington’s 1996 words) were far more similar than what Eurocentric authors used to think, they tend to recognise that substantial differences existed regarding political structures. For instance, Rosenthal & Wong (2011) acknowledge that Northwestern Europe and Eastern China were already on different institutional trajectories for several centuries before the Industrial Revolution. Similarly, Pomeranz (2011) accepts that his early claim of China’s Yangtze delta being on par with England as late as 1800 was exaggerated and now sees an earlier date between 1700 and 1750 as more realistic.

The most recent GDP per capita estimates indicate that England was no wealthier than other Western European countries such as Sweden and France before 1650; however, by this date the countries of Southern and Central Europe such as Portugal, Spain, and the German states were already at a lower level of development than England (Figure 2a). Moreover, Belgium, the central and northern states of Italy, and the Netherlands were wealthier than England before 1680, 1750, and 1805, respectively (Figure 2b). Hence, it seems clear that a Little Divergence between Western Europe on the one hand, and Southern and Central Europe on the other, occurred around 1350–1400. Given their importance for the Great Divergence debate, the causes of the Little Divergence within Europe will also be assessed in the following of the present article. Between 1350 and 1850 (at least), Japan, India, the Ottoman Empire, and Egypt were clearly more similar to South and Continental Europe than Western Europe; China as a whole seemed to lie in between the two groups (Figure 2c). Considering the high uncertainties that surround GDP per capita estimates of the past, it can be hypothesised that the wealthiest part of China (i.e., the Yangtze delta) was indeed on a par with England circa 1700. Compared to England, Italy and the Netherlands were on

---

10 Moreover, van Zanden (2008) adds that India and Indonesia had comparatively poor performances both regarding “horizontal” and “vertical” institutions.

11 Due to a lack of space, the limitations and uncertainties of the GDP estimates presented in Figures 1 and 2 are not discussed. The reader is invited to investigate these subtleties on its own, in particular choices regarding prices’ indexes and territorial changes over time in the different studies used to produce Figure 2. Moreover, it is important to note that linear interpolations have been performed whenever necessary to facilitate the graphical comparison of estimates.

12 The term Little Divergence was originally coined by Gorski (2005) to characterise the same intra-European bifurcation we talk about here. However, several authors also use this term to describe the diverging trajectories of China and Japan from 1850 onward (Dong et al., 2015, Sng & Moriguchi, 2014).

13 The most recent estimates of GDP per capita that exist for Central and South America concern Mexico and Peru from 1595 onward (Bolt et al., 2018). These estimates suggest that Mexico’s average GDP per capita was about as high as that of Spain, whereas that of Peru was never more than 90 percent of that of Spain.
higher but declining and stagnating paths respectively around 1700–1750, to the point that the Golden Age of the Dutch Republic should definitely be seen as the highest example of pre-modern efflorescent growth rather than the first instance of modern economic growth. Accordingly, 1700–1750 could be considered as the beginning of the Great Divergence if one considers that this phenomenon describes the new path towards modern sustained growth taken first by England. Instead of GDP per capita, scholars sometimes prefer to rely on real wages of skilled and unskilled construction workers in urban areas to compare standards of living in the late medieval and early modern eras. Angeles (2008) discusses the conditions under which the evolutions of real wages and GDP per capita are not aligned (e.g., change in occupational structure, income distribution, factor-specific biased technical change), but it is clear that the most recent data on real wages (Allen, 2011, Allen et al., 2011, Malanima, 2013) agree with the respective timings of the Little and Great Divergence that we have given.
above. However, we will return later to the important issue of real wages in the context of factor prices and directed technical change.

A legitimate criticism to the selection of 1700–1750 as the starting point for the World Great Divergence, and 1350–1400 as the beginning of the European Little Divergence, is that relevant differences between societies logically emerged before those disparities actually began to create significant discrepancies in wealth (in whatever way measured). This is true to such an extent that the origin of both the Little and Great Divergence could be traced back to the late Medieval period, or even long before. The investigation of the determinants of the Little and Great Divergence in the next section should helps us clarify this issue.

3. A review of the deep determinants of economic development

This review shows that different deep determinants of economic development are often connected by complex relationships. In particular, a strict distinction between cultural and institutional factors appears untenable nowadays, and actually, following such a distinction in the context of the Great Divergence debate would be counter-productive. Hence, as synthesised in Table 1, this section is organised around three major themes, namely 1. biogeography (Section 3.1), 2. culture-institutions (Section 3.2), and 3. contingency-conjuncture (Section 3.3). A given theory or study generally focuses on several determinants, so as with any classification, the allocation of a given theory or study to a particular category is often quite arbitrary. Moreover, some arguments concern the occurrence of the Great Divergence, others regard the persistence of this phenomenon, and the majority of deep determinants relate to both.

3.1. Biogeographical determinants

Biogeographical determinants can be classified around three main variants, namely 1. the long-lasting effect of local climate (temperature, humidity, rainfall), disease prevalence, and land suitability for agriculture; 2. the incidence of natural endowments in terms of sea access and overall openness of continents; and 3. the timing of the transition from foraging to farming (i.e., the agricultural Neolithic revolution) and its long-lasting impact on technical, cultural, and institutional developments.

3.1.1. Local climate, diseases, and land productivity

Kamarck (1976, p. 11) was one of the first to convincingly argue that climatic factors have hampered economic development in less developed countries of modern days through their impact on 1. agriculture (directly or through the diseases and pests afflicting animals and plants), 2. mineral discovery, and 3. prevalence of human diseases. Bloom & Sachs (1998) detail these points and argue that in Africa, in particular, tropical agriculture is faced with chronic problems of low yields and fragility due to low photosynthetic potential, high evapotranspiration, low and variable rainfall, highly weathered soils, veterinary diseases, and plant and animal pests. For these authors, evidence suggests that the burden of infectious disease (particularly malaria) is vastly higher in the tropics than in the temperate zones.

---

<table>
<thead>
<tr>
<th>Deep determinant of economic development</th>
<th>Most representative references</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biogeographical determinants (Section 3.1)</strong></td>
<td></td>
</tr>
<tr>
<td>Sea access and continental openness</td>
<td>Braudel (1996), Bloom &amp; Sachs (1998), Gallup et al. (1999)</td>
</tr>
<tr>
<td><strong>Cultural–institutional determinants (subsection 3.2)</strong></td>
<td></td>
</tr>
<tr>
<td>Ethnolinguistic fractionalisation and genetic diversity</td>
<td>Alesina et al. (2003), Ashraf &amp; Galor (2013), Spolaore &amp; Wacziarg (2013)</td>
</tr>
<tr>
<td>The rise (or not) of modern science and its possible religious roots</td>
<td>Lipsey et al. (2005), Mokyr (2011), Squicciarini &amp; Voigtländer (2015)</td>
</tr>
<tr>
<td><strong>Contingent–conjunctural determinants (subsection 3.3)</strong></td>
<td></td>
</tr>
</tbody>
</table>
Alsan (2015) advances and tests the hypothesis that the Tsetse fly (unique to Africa and transmitting the parasite *trypanosomiasis* which is harmful to humans and lethal to livestock) reduced the historical ability of Africans to generate an agricultural surplus. The author examines variation across ethnic groups within Africa and showed that by the end of the nineteenth century ethnic groups with climates more suitable for the Tsetse fly 1. had a higher participation rate of women in agriculture because they were less likely to use the plough; 2. had a higher propensity to use slaves because they were less likely to use draft animals; and 3. had lower population densities, fewer urban centres, and less centralised states because these regions were associated with shifting cultivation rather than more intensive agricultural techniques. Overall, Alsan’s (2015) findings provide strong evidence that the Tsetse fly inhibited the development of intensive agriculture using draft animals and the plough; this resulted in lower populations, less urbanisation, and less state development.

Michalopoulos (2012) provides evidence that natural heterogeneity in agricultural land quality and terrain slope is positively correlated with contemporary ethnic diversity, which is thought to be an important determinant of economic development today due to its effect on mutual trust as we will detail in Section 3.2. Interestingly, Michalopoulos (2012) shows that the influence of geographic diversity on linguistic diversity took place before 1500 because this relationship is not statistically significant in the parts of the world that witness significant population changes after 1500 (due to death and voluntary or involuntary migration). In the same vein, but focusing on Europe between 1500 and 2000, Buggle & Durante (2017) report that regions with higher inter-annual variability in precipitation and temperature were more closely connected to Medieval trade networks, indicating a higher propensity to engage in inter-community exchange because of higher levels of trust. Buggle & Durante (2017) also find that these regions were more likely to adopt participatory political institutions earlier on and are still characterised by a higher quality of local governments, all of which are said to be conducive to growth as we will see in Section 3.2.

Bentzen et al. (2017) empirically support the hypothesis advanced by Wittfogel (1957) that the large-scale investment and coordination necessary for irrigation also promoted strong authoritarian leadership and autocratic institutions that are detrimental to growth (see Section 3.2). To rule out reverse causality (authoritarian leaders may indeed be more willing and capable of undertaking large-scale irrigation works), and due to the lack of consistent historical irrigation data, the authors use a measure of irrigation potential based on exogenous geographical and climatic variables to identify societies that were based on irrigation historically. Then, Bentzen et al. (2017) show that countries with a high irrigation potential are less likely to be democratic nowadays, and they further show that land-ownership inequality is the most likely channel to explain this relation between natural suitability for irrigation and autocracy.

3.1.2. Sea access and continental openness


Regarding empirical evidence, Bloom & Sachs (1998) support econometrically that the failure of Africa to control diseases is not primarily the product of poor public health...
measures, unresponsive governments, or poverty but is rather due to the natural environment. Finally, to explain the long-term development lag of Africa, these authors point to five remarkable disadvantages in transport costs: 1. a great distance from major world markets in the northern mid-latitudes, in particular the separation from Europe by the Sahara desert; 2. a short coastline relative to the land area; 3. few natural coastal ports; 4. the highest proportion of landlocked states, and the largest proportion of the population within landlocked states, of any continent; and 5. the absence of rivers leading into the interior of the mainland that are navigable by ocean-going vessels (as are the Rhine, the Mississippi, the Amazon, and the Yangtze on other continents).

The statistically significant impact of geographical endowment (through climate and land openness) on per capita GDP growth is even more consistently demonstrated by Gallup et al. (1999). These authors conclude that sub-Saharan Africa is especially hindered by its tropical location, the high prevalence of malaria, and a low population density near the coast. By contrast Europe, North America, and East Asia have more favourable conditions on all four counts according to these authors.

3.1.3. Timing of the agricultural revolution

According to Diamond (1997, pp. 67–81), if Western Europe rules (for now) it is thanks to advantages in technology (guns, large sail ships), health (higher diseases resistance), and institutions (large markets, political organisation, property right protection) that we already present or on the verge to be established circa 1500 CE. Those advantages explain that Westerners colonised the New World, not the other way around, which implicates that they were then the first to industrialise. The author then asserts that if Westerners had such large technological and institutional advantages circa 1500, it is because Western Eurasia was the first world region to benefit from the early establishment of cities, writing, high population densities, and associated non-food producing elites that created and organised knowledge. For Diamond (1997, pp. 85–103), the ultimate cause of this fact is that Western Eurasia was the first world region to experience a transition from foraging to farming (called the agricultural Neolithic revolution). With several millennia of advance compared to other continents, domesticated plants and animals offered Western Eurasia a reliable source of food with high nutritional value, but also fertilisation, wool, leather, transport, ploughing, and military power that could feed a much greater population per unit area and consequently sustain a technology-inventing population. Diamond (1997, pp. 195–214) agrees with Crosby (2004, pp. 195–216) that the close physical proximity of man and animal also gave Eurasian inhabitants a relatively high resistance to animal-related germs such as diphtheria, smallpox, measles, and tuberculosis. The absence of an equivalent resistance to animal-related germs in the Americas proved to be devastating to native peoples during the colonisation of the New World since germs brought from Europe killed more native Americans than guns and swords.

As Diamond (1997, pp. 131–175) argues, if agriculture first emerged in the Hilly Flanks of Southwest Asia and spread to the rest of Eurasia, it is not because their inhabitants were more intelligent and better adapted to their environment; it is rather because their environment offered them a higher number of suitable plants and animals for domestication. Moreover, Diamond (1997, pp. 176–191) argues that the East–West orientation of Eurasia compared to the North–South orientation of America and Africa implies far less latitudinal variation (in terms of day length, seasonal variability, regimes of temperature and rainfall, and diseases) in Eurasia compared to other continents. Therefore, the spread of species, best-practices, and more generally technologies in the critical areas of agriculture and health was easier among
the relatively similar (climatic and ecological) regions of Eurasia compared to the regions of Oceania, Africa, and the Americas (that were moreover endowed with a higher number of natural barriers such as deserts, dense equatorial and tropical forests, and terrestrial bottlenecks such as the Isthmus of Panama).

Several econometric studies support Diamond’s (1997) claims. Olsson & Hibbs (2005) indicated that the unequal distribution of domesticable plants and animals accounts for around two-thirds of the regional variation in the estimated dates of the agriculture onset. The authors then showed that exogenous geography (continental size and axis, climate, latitude) and initial biogeographical conditions (number of domesticable plants and animals) account for 52% of the sixty-fold difference in contemporary per capita income observed in a broad international cross-section of 112 countries. Putterman (2008) argues that a potential limitation of Olsson & Hibbs’s (2005) study is that the estimate for the date of Mesopotamia’s agricultural transition is used not just for present-day Iraq and Jordan but also for Great Britain and Europe as a whole; whereas archaeological evidence suggests that agriculture was not established in Great Britain until some 5000 years later than modern day Iraq. However, using a new set of estimates of the year of transition to agriculture in a large number of countries, Putterman (2008) showed that the time elapsed since the Neolithic Revolution explains up to 34% of the variance in estimated income in 1500. Putterman & Weil (2010) demonstrated that adjusting the time of the transition to agriculture so that it reflects the history of a population’s ancestors rather than the history of their present location greatly improves the accuracy of those indicators for predicting contemporary GDP levels. In the same vein, Ashraf & Galor (2011) found a highly statistically significant positive effect of regional differences in land productivity and the number of years elapsed since the Neolithic revolution on local population density (which is arguably a better estimate of prosperity) in the years 1 CE, 1000 CE, and 1500 CE. Similarly, Ang (2015) shows econometrically that countries that experienced earlier transitions to agriculture were subsequently more capable of adopting new technologies in 1000 BCE and, to a lesser extent, in 1 CE and 1500 CE.

In summary, the most recent empirical evidence seem to indicate that Diamond’s (1997) thesis on the long-lasting effect of the Neolithic revolution (and to a lesser extent perhaps the overall impacts of biogeographical factors) offers compelling explanations for intercontinental differential developments up to 1000–1500, but it cannot adequately explain the current level of economic inequality between countries of the same world region. Hence, as already seen with Michalopoulos (2012) and Bentzen et al. (2017), it seems that if the endowment of biogeographical factors has surely had a significant direct effect on the development of world regions up to 1000–1500, the impact of geography on the contemporary comparative development of nations has then been mostly channelled indirectly through cultural and institutional changes.

However, after controlling for the effects of a large number of geographical factors (including absolute latitude, access to waterways, distance to the technology frontier, and the share of land in tropical versus temperate climatic zones, which may have had an impact on aggregate productivity either directly by affecting the productivity of land, or indirectly via the prevalence of trade and the diffusion of technologies), Ashraf & Galor (2011) assert that the effects of land productivity and the number of years elapsed since the Neolithic Revolution on the per capita income estimates of 1500 CE are not significantly different from zero, which contradicts the results of Olsson & Hibbs (2005), Putterman (2008), and Putterman & Weil (2010) on this point.
3.2. Cultural–institutional determinants

Cultural–institutional determinants of long-term economic development can be organised into three categories, namely 1. the long-lasting influence of ethnolinguistic fractionalisation, possibly in relation to genetic diversity; 2. the effect of religions on human capital accumulation and institutional divergence; and 3. the impact of religion on the rise (or not) of modern science as a pre-requisite for an industrial revolution. Before reviewing the evidence for these different deep determinants, we need to define the notions of culture and institutions and argue for the necessity to analyse these determinants together rather than separately.

3.2.1. Culture, institutions, and their co-evolution

According to Guiso et al. (2006, p. 23) and Mokyr (2017, p. 8), culture is a set of beliefs, values, and preferences, capable of affecting behaviours, that are socially (not genetically) transmitted fairly unchanged from generation to generation and that are shared by some subsets of society (ethnic, religious, and social groups). The problem with a cultural-only approach lies in the difficulty of establishing a straightforward causal link between core beliefs and preferences on the one hand, and economic performance on the other. The main channel through which culture affects economic development is trust, defined by Gambetta (1988, p. 217) as “the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action”.16 Several econometric studies have indeed shown that trust and civic cooperation have a significant positive correlation with aggregate economic activity (see Algan & Cahuc, 2013 for a recent review). However, despite the various impacts that cultural traits can have on growth and the significant path-dependency of culture (i.e., the fact that culture is a historical heritage), the endogenous nature of culture also implies that economic development is associated with shifts toward values of rationality, tolerance, trust, and participation (Inglehart & Baker, 2000). This is one reason that could prevent establishing a direct causal link between culture and economic growth. Another reason is the importance of institutions and their link with cultural traits.

Following North (1990) and Acemoglu & Robinson (2012), institutions correspond to the formal mechanisms—constitutions, laws, organisations’ rules—and their enforcement characteristics, through which social choices are determined and implemented.17 In combination with the distribution of economic resources, political institutions determine the distribution of political power across different socioeconomic groups; these in turn shape economic institutions that direct economic performance and resource allocations. This synergistic relation between economic and political institutions is enriched by the distinction between inclusive and extractive institutions (see Acemoglu & Robinson, 2012, pp. 74–81).18 The

16 A complementary definition of generalised trust comes from Coleman (1990), according to whom individuals trust if they voluntarily place resources at the disposal of another party without any legal commitment from the latter, but with the expectation that the act of trust will pay off.

17 On the contrary, informal constraints such as norms of behaviour, convention, and self-imposed codes of conduct, “come from socially transmitted information and are part of the heritage that we call culture” (North, 1990, p. 37).

18 Examples of inclusive economic institutions include secure private property, an unbiased system of law, and a provision of public services that provide a level playing field for markets that permit the entry of new businesses and allow people to choose their careers (Acemoglu & Robinson, 2012, pp. 74–75). Extractive economic institutions have opposite properties and are designed to extract incomes and wealth from one subset of society to
central idea of the institutional hypothesis is that economic growth and prosperity are associated with inclusive economic and political institutions, while extractive institutions typically lead to stagnation and poverty.\footnote{The typical modern example of a rather rapid divergence in economic performances following different institutional choices concerns the Democratic People’s Republic of (North) Korea and the Republic of (South) Korea (see Acemoglu & Robinson, 2012, pp. 70–73).} Hence, for proponents of the institutional hypothesis, the occurrence and persistence of the Great Divergence mainly come from the fact that some nations managed to develop inclusive institutions that fostered economic development, whereas others did not.

If numerous scholars have claimed that culture does not directly affect economic growth, but instead plays an indirect role through institutions (Greif, 1994, 2006, Todd, 2019, Zerbe & Anderson, 2011), an equal number of researchers have argued to the contrary that institutions shape cultural traits (Alesina & Fuchs-Schündeln, 2007, Becker et al., 2016a, Grosjean, 2011). The most recent theoretical frameworks and empirical evidence now suggest that culture and institutions are connected through a co-evolutionary relation that is so tight that trying to analyse them separately would be misleading and counter-productive (Aghion et al., 2010, Alesina & Giuliano, 2014, 2015, Bisin et al., 2018, Guiso et al., 2015, Lowes et al., 2017, Tabellini, 2008, 2010).

3.2.2. Ethnolinguistic fractionalisation and genetic diversity

A significant number of studies examine the impact of ethnic, linguistic, and religious fractionalisation on modern economic growth. These studies are potentially important for explaining the persistence of the Great Divergence but not its occurrence. Knack & Keefer (1997) have shown that contemporary levels of trust and norms of civic cooperation (that positively affect modern economic growth) are stronger in countries that are less polarised along class or ethnic lines. Similarly, Easterly & Levine (1997) assert that modern cross-country differences in ethnic diversity are positively correlated with domestic variations in public policies, political instability, and other economic factors associated with long-run growth. Alesina et al. (2003) add nuance to these results and explain that ethnic and linguistic fractionalisation are likely to be important determinants of economic success but that strong correlation with other variables (geographical ones in particular) greatly complicates the evaluation of the size of these effects. Finally, Alesina & La Ferrara (2005) argue that contemporary fractionalisation has adverse consequences on growth and productivity only in nondemocratic regimes, while democracies do not have such issues.

A recent set of publications goes deeper to explain both the occurrence and persistence of the Great Divergence through culture. They explore the long-lasting influence of the genetic composition of populations on their cultural characteristics and their consequences on the comparative economic performances of societies. The first formal theory for the causal effect of genetic material on differential economic performance was advanced by Clark (2007). This scholar proposes that Darwinian natural selection of the fittest (in his view the richest) endowed with growth-compatible characteristics (entrepreneurial and hard-working spirits, patience, and propensity to innovate) can explain the phenomenon of the Great Divergence, on the unexplained premise that such a natural selection mechanism was more

benefit a different subgroup (Ibid., p. 76). Inclusive political institutions are those that are sufficiently centralised and pluralistic, and they become extractive when either of these conditions fails (Ibid., p. 81).
active in England than in the rest of the world during the centuries preceding the Industrial Revolution.20

With a quite different approach, Spolaore & Wacziarg (2009) assert that “genetic distance”, a measure associated with the time elapsed since the last common ancestor of two populations, has a statistically significant effect on income differences across countries (even controlling for measures of geographical distance, climatic differences, transportation costs, and measures of historical, religious, and linguistic distance). The authors interpreted these findings in terms of barriers to the diffusion of development from the world technology frontier. Spolaore & Wacziarg (2009) found that the effect of a country’s relative genetic distance to the world technology frontier (Britain in the nineteenth century and the United States in the twentieth century) on their economic distance is positive and larger than the effect of absolute genetic distance. In a subsequent study, Spolaore & Wacziarg (2014) show that genetic distance measured relative to the world’s technological frontier trumps absolute genetic distance as an explanation not only for bilateral income differences but also for the different usage of specific technological innovations. Furthermore, the authors documented that 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.

An alternative view is proposed by Ashraf & Galor (2013) who use data on genetic diversity from the 53 ethnic groups across the globe that constitute the Human Genome Diversity Cell Line Panel and show that migratory distance from East Africa has an adverse effect on genetic diversity so that genetic diversity is higher for natives of Africa, lower for natives of Asia, Oceania, and South America, and intermediate for natives of Europe. To the authors’ mind, genetic diversity is both negatively associated with the extent of cooperative behaviour as it raises the likelihood of mistrust and positively related to innovative activity as measured by the intensity of scientific knowledge creation. For Ashraf & Galor (2013) the degree of genetic diversity in a society may provide a wider spectrum of traits appropriate for developing advanced technological paradigms (possibility of expanding the society’s production frontier), but it may also reduce trust, cooperation, and hence the efficiency of the production process. In support of their theory, Ashraf & Galor (2013) observed a hump-shaped relationship (i.e., an inverted U curve) when population density in 1500 CE or the level of income per capita in 2000 CE (the best measure of development levels at these respective time according) are plotted as a function of genetic diversity with various control variables.

As could have been expected, these recent works on the relationship between the genetic composition of populations and comparative economic performance of societies have triggered a vibrant debate. Benjamin et al. (2012) were the first to provide a comprehensive reflection on the promises and pitfalls of this emerging field of research baptised “genoeconomics”. The work of Ashraf & Galor (2013) received extremely harsh criticisms from anthropologists and archaeologists (d’Alpoim Guedes et al., 2013) to which a response was given in an open letter available online (see Ashraf & Galor, 2014). Among dozens of reviews, the evolutionary theory of Clark (2007) has come in for particularly vigorous criticism from four referees (Grantham, 2008, McCloskey, 2008, Persson, 2008, Voth, 2008) to which Clark (2008) responded in the same journal. Hence, if the interaction between genetic and cultural evolution has been intensively explored since the 1980s (more recent references include Richerson & Boyd, 2005 and Jablonka & Lamb, 2014), additional research is needed to clarify

the complex relations existing between genetic and cultural inter-generational transmission of traits on the one hand, and economic outcomes on the other (see the complementary literature reviews of Spolaore & Wacziarg, 2013 and Collins et al., 2016).

3.2.3. Religions and cultural–institutional changes
Among contemporary scholars, Landes (1998) and Jones (2003) are usually cited as the main representatives of a Eurocentric explanation of the Great Divergence focusing on cultural differences. Landes’ (1998, p. 516) judges that “if we learn anything from the history of economic development, it is that culture makes all the difference. (Here Max Weber was right on.).” Landes refers to Weber’s (1930) theory stressing that the Protestant Reformation, and the Protestant work ethic it spurred in the sixteenth century, played a key role in the rise of a modern industrial society in Western Europe. Weber argues that, contrary to Catholicism, Protestantism defines and sanctions an ethic that is conducive to business success because the Protestant work ethic makes people work harder, work more efficiently, and is akin to entrepreneurship. According to Weber, these complementary social aspects should together be the main cause of the establishment of self-sustaining economic growth associated with capitalism in Western Europe. Focusing on the most recent econometric studies, Becker et al.’s (2016b) literature review is certainly one of the best synthesis on the causes and consequences of the Protestant Reformation. First, these authors recall that early criticisms of Weber’s theory made by Tawney (1926) and Samuelsson (1961) pointed out that capitalism had already taken firm root in Italy and the Netherlands prior to the Protestant Reformation, with complex financial systems that partly caused the past efflorescences of these countries. Moreover, there is reason to believe that key aspects of what would become the Protestant work ethic existed in certain territories long before the emergence of Protestantism (Akcomak et al., 2016; Andersen et al., 2017; Todd, 2019, chapter 5). Nevertheless, the most recent econometric studies support the ideas that today’s Protestants work more hours in Germany (Spenkuch, 2017), have less of a preference for leisure in Switzerland (Basten & Betz, 2013), and across all Europe suffer more displeasure when they are unemployed (van Hoorn & Maseland, 2013); this evidence is consistent with the idea that Protestants (want to) work hard. Moreover, because Luther wanted all Christians to read the Bible, Becker & Woessmann (2009) advance and empirically demonstrate that the Protestant Reformation fostered mass education and narrowed the gender gap in school enrolment, which was instrumental to see an increase in literacy rates and consequently in human capital accumulation.

However, econometric studies obtain mixed results when assessing the relationship between Protestantism on the one hand and the advent of industrial capitalism and sustained economic development on the other. Concentrating on European countries in the mid- to late-nineteenth century, Delacroix & Nielsen (2001) find a significant relationship between Protestantism and savings bank deposits per capita, supporting the idea of superior Protestant frugality; however, they do not find a relationship between the share of Protestants and

21 However, McCloskey (2016) is probably a serious contender for such a title with her idea of bourgeois virtues at the origin of liberalism and the English take-off (with however, an underlying role for four major events, namely the German Reformation, the Dutch Revolt, the American and French Revolution, and the Scottish and Scandinavian early mass literacy).

22 Nevertheless, it would be unfair to say that Landes (1998) thinks that culture alone explains all the differences among countries' abilities to generate wealth. In his own words, “economic analysis cherishes the illusion that one good reason should be enough, but the determinants of complex processes are invariably plural and interrelated” (Ibid., p. 517).
wealth per capita or other indicators of economic activity (date of founding of the principal stock exchange, extension of the railroads network in 1870, male labor force in industry). Furthermore, Delacroix & Nielsen (2001) point to countries with significant Protestant populations that were notable laggards in capitalist development (such as the Nordic countries) and countries with a majority of Catholic inhabitants that were early developers (such as Belgium and France). Alternatively, Cantoni (2015) uses city growth (for 272 entities in the Holy Roman Empire from 1300 to 1900) as a proxy measure of pre-modern economic growth, and he finds no effects of Protestantism on economic growth in a generalised differences-in-differences set-up.

Following the previous insight of several sociologists and political scientists, Arruñada (2010) suggests and empirically supports that Protestantism is conducive to capitalist economic development; however, this is not because of the direct psychological Weberian work ethic route but rather by promoting an alternative social ethic (e.g., tolerance towards foreigners) and a higher willingness than Catholicism to support a secularisation of political institutions that facilitates impersonal trade. Consistent with this idea, Rubin (2017, pp. 149–168) argues that after the Reformation, Protestant rulers could not depend on the religious elite for legitimacy and thus turned to parliaments to support their rule and provide them revenue. In line with Acemoglu et al. (2005) to whom we will return later, Rubin (2017, pp. 149–168) makes the case that, as inclusive institutions, parliaments generally had interests more aligned with long-run economic development than the religious elite. According to Rubin (2017, pp. 149–168), this helps explain why England and the Dutch Republic experienced substantial economic growth soon after their Reformations, while Catholic Spain and the Muslim Ottoman Empire lagged behind in spite of their military and territorial dominance. Kuran (2012, 2016) provides a complementary argument by asserting that Islamic economic institutions, which had benefited Middle Eastern economies in the early centuries of Islam, then acted as a drag on development by slowing or blocking the emergence of central features of modern economic life. Those extractive economic institutions include the following: 1. Islam’s original tax system, which failed to protect property rights; 2. the waqf (a non-state institution ruling property donation to the civil society), whose rigidity hampered the development of civil society; and 3. private commercial enterprises, whose small scales and short lives blocked the development of private coalitions able to bargain with the state. According to Kuran (2012, 2016), these extractive institutions contributed to features that sustained autocracies and kept democracies unstable (high corruption, low trust, widespread nepotism, and a high tolerance for law-breaking), which has had long-lasting effects on the economic development of the Middle East to the present day.

3.2.4. The rise (or not) of modern science and its possible religious roots

Several scholars such as Goldstone (2009), Jacob (2014), and Mokyr (2011, 2017) attribute much of the credit for the burst of innovations and accelerated diffusion of best practices after 1750, not to mass education in general, but to the scientific culture that emerged with

---

23 Delacroix & Nielsen (2001) thus conclude that the popular success of Weber’s theory on the important role of Protestantism for the early development of industrial capitalism is mainly derived from selected anecdotal evidence, fortified, through retrospective imputation, by the perceived well-being of contemporary Protestant countries.

24 Consistent with Kuran (2012, 2016), Shatzmiller’s (2011) examinations of the role of Islamic institutions between c. 750 and c. 1100 show that economic growth was visible in the key indicators of the Caliphate economy. Hence, there is indeed nothing intrinsic to Islamic institutions that impaired economic growth.
the European Enlightenment. They argue that Western European societies were particularly
dynamic and inclined to see technical breakthroughs in the eighteenth century, thanks to the
increase over the previous 200 years in the number of printed books, publishers, scientific
societies, university networks, relatively accessible public lectures, and growing day-to-day
exchanges between scientists, engineers, and craftsmen. It is important to understand that
these scholars do not denigrate the many scientific breakthroughs that episodically originated
in China and in Islamic countries. They rather highlight the earliness of Great Britain in cre-
ating an intellectual, social, and institutional environment able to transpose useful knowledge
(i.e., ideas and inventions that often came from other European countries such as France,
Belgium, the Netherlands, or Germany), into applied knowledge (i.e., skills and technical change).
We will see in Section 3.3.4 that, as Crafts (2011) argues, the “factors-prices-
incentive” explanation of Allen (2009, 2011) for the rise of technology is complementary
rather than alternative to the “scientific-culture-oriented” argument emphasised here.

Charting manuscripts and then printed books production from the sixth to the eighteenth
centuries, Buringh & van Zanden (2009) provide empirical evidence for the close association
between books production and regional patterns of economic growth and decline from the	twelfth century onward. Furthermore, they show that monasteries dominated books
production from the sixth to the tenth centuries, whereas demand from cities and universities
elites drove the continuous growth of the book industry from the eleventh century onward.
Buringh & van Zanden (2009) note that the growing literacy of the urban population, the
long-term increase in their incomes, and rapid technical change in the production of books
(in particular after the invention of Gutenberg’s movable printing press in 1454) dominated
the book diffusion process in the early modern period.25 Moreover, Squicciarini & Voigtländer
(2015) provide direct evidence for the importance of useful knowledge for industrialisation.
As a proxy for scientific elites, the authors use Encyclopédie subscriber density and show that
this measure of “upper-tail knowledge” is strongly associated with other indicators of local
scientific activity, both before and after the Encyclopédie was printed in the mid-eighteenth
century. Squicciarini & Voigtländer (2015) then show that upper-tail knowledge is a strong
predictor of city growth after the onset of French industrialisation. Furthermore, by joining
data on British patents with a large French firm survey from the 1840s, it appears that
scientific elites caused productivity increases in innovative industrial technology which were
then associated with economic growth. Moreover, Squicciarini & Voigtländer (2015) showed
that literacy levels representing human capital of the general population are associated with
development in the cross-section, but they do not predict growth.

However, Khan (2018) argues to the contrary that elites and allegedly upper-tail knowledge
were neither necessary nor sufficient for technological productivity and economic progress
before and during the first part of the British Industrial Revolution. According to Khan
(2018), the evidence from the education and patenting of the great inventors in Britain

25 Angeles (2017) argues that 1. China’s preference for block printing technology over movable type was rational
given its complex logographic writing system and that 2. this choice did not lead to larger printing costs which
could explain China’s modest level of book production when compared to Europe’s. A purely political reason
such as the one advanced by Cosgel et al. (2012) in the Ottoman case may better explain China’s delay in book
production in the pre-modern period. For Cosgel et al. (2012), Ottomans rulers regulated the printing press
during almost three centuries to prevent the loss it would have caused to their net revenue by undermining the
legitimacy provided by religious authorities. Cosgel et al. (2012) add that European rulers had little reason to stop
the spread of printing because the legitimising relationship between European religious and political authorities
was undermined over a century prior to the invention of the press.
between 1750 and 1930 indicates that scientists, engineers, or technicians were not well represented, and their contributions remained unspecialised until the late nineteenth century. Hence, Khan (2018) suggests that the formal acquisition of human capital and production of useful knowledge did not play a central role in generating new inventions. For him, the acquisition of applied knowledge depended more on transmission mechanisms from fathers to children within households, and mostly from masters to apprentices within crafts and guilds. A potential consensus between these mixed results is offered by de Pleijt et al. (2019) who performed a quantitative assessment of the effect of industrialisation, captured by the number of steam engines per person installed in England by 1800, on the average working skills of 2.6 million workers. They found support for a causal relationship between the diffusion of steam engines and demand for higher skills. Furthermore, de Pleijt et al. (2019) show that early industrialisation was negatively associated with primary schooling and with the acquisition of literacy skills for women. Overall, de Pleijt et al.'s (2019) findings tend to confirm Mokyr’s (2011) conclusion that basic education and the associated human capital was not a key ingredient in England’s early industrialisation; and 2. show that the causal relationship running from useful knowledge (i.e., scientific ideas and inventions) to industrialisation highlighted by Squicciarini & Voigtländer (2015) was complemented by a causal relationship going from industrialisation to applied knowledge (i.e., skills).

While the importance of useful scientific knowledge for the Industrial Revolution is still being debated, so is the origin of modern science through the European Enlightenment. For Lipsey et al. (2005, pp. 225–289) the roots of modern science lie in the second half of the medieval period in Western Europe which saw the development of pluralistic societies that ultimately freed natural philosophers to seek explanations of the world in mechanical laws. These authors also assert that the absence of an early economic takeoff in China and Islamic countries is explained by the failure of these countries to develop anything similar to modern science because of inappropriate institutions. Lipsey et al. (2005, pp. 225–289) think that despite its high level of bureaucratisation, the inability of China to develop an independent version of modern science has more to do with the absence of repositories for saving and organising cumulative knowledge, whereas Europe institutionalised scientific research through its universities and scientific societies. Furthermore, Lipsey et al. (2005, pp. 225–289) argue that Islam is an occasionalist doctrine in which the state of the world at any moment in time is contingent on the particular will of God. On the contrary, the doctrine of Christian naturalism posits that God created the world according to natural laws and then endowed humans with free will to determine their own affairs. For Lipsey et al. (2005, pp. 225–289), this difference was decisive for the development of science in early modern Europe whereas Islam was relatively hostile to free inquiry and mechanistic science. Chaney (2019) provides a complementary hypothesis to this last idea. He argues that, from the eighth to the twelfth centuries, the Islamic world’s Golden Age of scientific production was closely linked to the region’s secular bureaucratic structures. According to Chaney (2019), these institutions encouraged scientific production by both increasing demand for scientific output and by constraining religious leaders opposed to the study of scientific topics. The use of slave soldiers and the eventual militarisation of government led to the collapse of this bureaucracy and the empowerment of religious leaders in the twelfth century. In addition to taking charge of judicial administration, local police, irrigation, public works, and taxation,

26 See Ogilvie (2014) for an overview of the debate on the positive and negative impacts of guilds on economic development.
The religious elites worked to restrict scientific production because they believed that the unrestricted study of science would undermine their societal position and control over the population. Consistent with Shatzmiller (2011), Chaney (2019) does not believe that Islam is uniquely inimical to scientific development. Instead, in line with Rubin (2017), the evidence presented by Chaney (2019) support that secular state structures can play a fundamental role in constraining rent-seeking religious elites aiming to restrict scientific research.

Finally, Baechler (2002, pp. 239–240) notes that all great civilisations have been defined and imbued with a great religious doctrine. Therefore, it seems difficult to explain the emergence of modern science and the transition to sustained growth solely as a result of Christianity. A given religion has no intrinsic and determined impact on knowledge accumulation and economic potential. Rather, the long-run economic effects of a religious doctrine are only understandable when considered in conjunction with their associated institutional changes. Through institutional changes, religious doctrines indeed affect a host of variables that are important for economic growth, including human capital, governance, entrepreneurship, social ethic, and social networks. Consistent with this idea, Goldstone (2012), O’Brien (2013), and Jin (2015) argue that the emergence of modern science in Europe is also due to a series of events (namely inter-state competition, the recovering of Greek knowledge and the Renaissance, the discovery of the Americas which shattered faith in Ptolemy’s Europe-centred geography and Earth-centred astronomy, the Protestant Reformation, the Copernican model of the universe and its refinement by Kepler and Galileo, the separation between church and state, the legal revolution, the printing revolution, and a variety of other technical discoveries) that together made it increasingly hard to maintain the fundamental doctrines and philosophy of Western Christianity. For these three scholars, it is thus a unique European combination of factors that led to the emergence of a widespread systematic scepticism towards existing knowledge and a push towards the modern scientific method (characterised by the use of specific instruments, language, quantification, and mathematisation of artificially created phenomena, and the enforcement of rules of objectivity, falsifiability, and replicability).

If contingent events and the global conjuncture that combined them have had such an impact on the emergence of modern science, it is legitimate to wonder how they have also affected other cultural aspects, institutions, and the development trajectories of nations.

### 3.3. Contingent–conjunctural determinants

To obtain a more accurate explanation of the early Western European (and first of all British) take-off, cultural–institutional determinants of long-term growth must also be characterised with several accidental events that worked in conjunction from the Middle Ages to the eighteenth century, namely: 1. the advent of the Black Death in the middle of the fourteenth century and its consequences, especially on the European marriage pattern; 2. the technological and institutional divergences resulting from different state capacities between the divided European nation-states, the unified Chinese empire, and the slavery-dependent

27 Indeed, if religions probably played a role to explain the European location of the rise of modern science, empirical evidence provided by Bénabou et al. (2015) suggest that there is a significant and robust negative relationship between contemporary levels of religiosity and patents per capita in both international and US cross-state data.

28 For example, the recent literature on the Protestant Reformation and economic outcomes documents to a wide range of mechanisms. Where the Reformation took hold, it fundamentally altered political, legal, and social institutions, resulting in the ascendancy of parliaments, the secularisation of law, increased emphasis on education, and the precursors of the welfare state (Becker et al., 2016b, pp. 21–22).
Islamic sultanates and empires of the Middle East; 3. the relative sizes of the Atlantic and Pacific oceans, the consequent occurrence of the Atlantic slave trade and its impact on the institutional divergence of Western European economies and their colonies; and 4. the crucial role of the uneven global distribution of coal and its transformation into mechanical energy by the steam engine.

3.3.1. The Black Death and the European marriage pattern

The Black Death is the paradigmatic example of an exogenous shock causing major and long-lasting divergence across regions. It killed around one-third of the population in Europe from its first appearance in 1348 to the end of the fourteenth century. Repeated outbreaks of the plague remained the strongest obstacle to population growth in most of Europe for the next three centuries. Population returned to 1300 levels during the fifteenth century in the Low Countries, during the sixteenth century in France, Spain, and Italy, and only in the eighteenth century in England and Wales. The population of Egypt likely did not return to pre-Black Death levels until the nineteenth century (Pamuk, 2007, pp. 293–294). Building on previous scholars, Pamuk (2007, p. 294) reported that the direct consequence of the overall labor force decline was a fall in total output that was not as large as the population decline, so output per capita increased in the 100 years following the first occurrence of the plague. At the same time, the large decreases in population and the labor force resulted in a doubling of urban real wages in most countries and cities of Europe and around the Mediterranean. Land became more abundant relative to labor, so land rents and interest rates went down in both absolute terms and relative to wages, and prices of agricultural goods declined relative to manufactured products. Labor scarcity generated an incentive for labor-saving technical change, and both agriculture and manufacturing began developing along more capital-intensive lines.

Eventually, mortality rates declined as the plague lost momentum, and population and output increases resumed in the middle of the fifteenth century. As population and output increased, wages and per capita incomes in both the urban and rural economies peaked around 1450 and then began to decline. However, not all regions reacted the same and real wages declined faster and longer in Southern and Continental Europe than in Northwestern Europe where they even stabilised during the second half of the sixteenth century (Pamuk, 2007, pp. 296–301). Overall, between 1350–1600, real wages of both skilled and unskilled workers in the leading cities of northwestern Europe (London, Oxford, Amsterdam, Antwerp) increased relative to the real wages in Northwestern, Southeastern, and Continental European cities (Florence, Valencia, Paris, Vienna, Munich, Krakow) which aligned with Mediterranean areas (Istanbul, Cairo). In the same time, urbanisation rates (a proxy for agriculture-to-industry shift and productivity gains) of Northwestern Europe countries caught up with urbanisation rates of territories in Southern Europe. For Pamuk (2007, pp. 309–311), the origin of this intra-European Little Divergence lies in the higher flexibility of northwestern institutions that was probably present before the Black Death but was plainly revealed only by this major exogenous shock. The flexibility of urban guilds (and to a lesser extent the land-ownership system) seems to have been particularly instrumental for explaining that Northwestern Europe responded better than Southern and Continental Europe to the new environment generated by the Black Death.

Another important outcome of the Black Death noted by Pamuk (2007, pp. 307–308) relates to changes in the demographic regime (i.e, the relationship between deaths, marriages, and births). The available evidence suggests that, in England and the Low Countries in particular, the shortage of workers after the Black Death and the availability of well-paid
employment attracted younger women into the labor force. As a result, a larger fraction of the households became dependent on wage labor in the two centuries following the Black Death. Even if debates still exist about the role of other determinants in this shift, the increase in employment opportunities for females appears to have led to a rise in age of marriage and a decline in fertility in both England and the Low Countries, a phenomenon called the Western European marriage pattern (EMP). The importance of the EMP for the Great Divergence between Europe and the rest of the world has been the subject of a long and intense debate, of which we only present the most recent opposing elements. Dennison & Ogilvie (2014) strictly define the EMP using its original characterisation by Hajnal (1965); these are a high age for women at first marriage, a high percentage of singles, and a high percentage of nuclear (i.e., non-complex) households. Based on 2,622 observations of female age at first marriage, 1,172 observations of female lifetime celibacy, and 911 observations of the kin complexity of households covering 39 European countries from 1500 to 1900, Dennison & Ogilvie (2014) found no evidence of a consistent relationship between the EMP and higher economic growth, nor any empirical evidence that the EMP improved female autonomy, increased human capital investment, enhanced demographic responsiveness to economic conditions, or created growth inducing cultural norms.

Criticising the analysis of Dennison & Ogilvie (2014), Carmichael et al. (2016) have a substantially different approach to the EMP concept that they embed into the larger ethnographic literature related to family systems (see Todd, 2019). Precisely, Carmichael et al. (2016) analyse the EMP as a dynamic family system characterised by two key principles; these are partners’ consensus to marry and neo-locality (i.e., the creation of a new household independent of both the groom’s and bride’s parents). By using such a broader definition of the EMP, a woman’s position is also strengthened by their right to own property, have a share in inheritance, and that marriage is strictly monogamous (even, de jure, for the elite) and exogamous (one marries outside the kimgroup, leaving more room for choice). The EMP, as Carmichael et al. (2016) now define it, is exceptionally “girl-friendly” and therefore looking solely at marriage ages is not very helpful in this perspective. Carmichael et al. (2016) assume a consistent link between their female-friendly EMP at the micro level and economic performance through the strengthened position of women because such an agency restrains population growth, enhances human capital formation of women and their offspring, and encourages their access to labor and capital markets. As stated again by Dennison & Ogilvie (2016) in their response to Carmichael et al.’s (2016) criticisms, it seems clear that the EMP, in its strict Harjnal-ish definition (to which colleagues of S. Carmichael used to comply in previous articles), did not consistently improve economic performance and thus cannot be considered as a cause of the Great Divergence. Despite Carmichael et al. (2016) continued use of the term EMP, the cultural–institutional interaction towards which their analysis has shifted is something else related more broadly to family systems and their classification.

3.3.2. Intra-European political fragmentation versus unified empires of the (far and middle) East
From the fourteenth to the nineteenth centuries, Europe had violent competing states, whereas Asian sub-continents (i.e., China, India, and the Middle East) were relatively more unified and centralised territories (Pomeranz, 2000, p. 194; Baechler, 2002, pp. 92–110).

Hajnal (1965) also mentions a large share of the population working as life-cycle servants and a low age gap between marriage partners.
This difference is thought to have had three main consequences on the technological and institutional divergence of Europe, Eastern and Southern Asia, and the Middle East.

Firstly, Europe’s internal military competition seems to have been instrumental for generating a comparative advantage in military technologies relative to other Eurasian regions, in particular regarding the mastery of gunpowder. Against this generally accepted claim, Andrade (2016) argues that China also experienced rapid military change and was probably on par with Europe c. 1700 in terms of terrestrial military technology and organisation. However, Andrade (2016) points out that the superiority of Europe over China in terms of naval matters was demonstrated during the First and Second Opium War (1839–1842 and 1856–1860, respectively). Andrade (2016) explicitly highlights that the West benefited from the application of science to military technologies, for example in ballistics. It was very difficult for China to catch up in this regard, not only because of its scientific and technological “backwardness”, but also because the nineteenth century Chinese state was weak and underfunded after a period of relative peace since the 1760s. For Hoffman (2015), incessant warfare among closed nation-states is not sufficient to explain the astonishingly rapid growth in Europe’s military sector from the Middle Ages. According to this author, Western Europeans rulers also had lower political costs of summoning resources (through taxes), higher incentives to not use older military technologies, and fewer obstacles to adopting military innovations compared to China, Japan, India, and the Ottoman Empire. Hence, for Hoffman (2015), it is political history that explains how Western Europe acquired an insurmountable lead in gunpowder technology. For him, this determined which states established colonial empires and ran the slave trade, and even which economies were the first to industrialise.

Secondly, Europe’s internal warfare implied fiscal innovation to raise revenue, which seems to have been instrumental for the emergence of state capacity in European countries from the mid-sixteenth century onward (Tilly, 1990). Hoffman & Rosenthal (1997, p. 36) report that, by modern standards, military expenditure was already large in peace time and loomed even bigger during the wars that almost constantly haunted pre-modern Europe. Karaman & Pamuk (2013) refine the link between warfare, fiscal capacity, and state building in Europe taking into account the economic structure (proxied by urbanisation rate) and political (authoritarian versus representative) regime of nations. They find that, under pressure of war, it was representative regimes in more urbanised-commercial economies and authoritarian regimes in more rural-agrarian economies that could best align domestic interests towards state-building. Voigtländer & Voth (2013) design a model in which the Black Death acts as a major shock that displaces the population dynamics of Europe on a completely different path than China. Because of Engel’s Law (consumers grown richer after the Black Death spent relatively less on food, and more on manufacturing goods), demand for urban products increased, and urban centres with unhealthy conditions grew in size, which further pushed

---

30 As recalled by Johnson & Koyama (2017, p. 2), state capacity describes the ability of a state 1. to garner enough tax revenues from the economy to implement its policies (fiscal capacity) and 2. to enforce law and order across the entirety of the territory it claims to rule (legal capacity). Johnson & Koyama (2017) also provide the most recent literature review on the different paths taken by European countries to build their state capacities and assess the consequences on economic development. The article suggests that England’s small size, but mostly its low ethnolinguistic diversity compared to Prussia, France, Spain, and the Habsburg territory, has been a significant advantage in consolidating its state capacity. High capacity states in Europe contributed to economic development through a range of channels: they made economies more robust to negative shocks such as warfare, helped to integrate domestic markets, and laid the foundations for relatively impartial bureaucracies and the rule of law.
up aggregate death rates. Moreover, since war is a “superior good” for princes in this model, larger tax revenues and higher war frequency went hand-in-hand. The consequence is an increase in state-capacity and urbanisation with a concomitant decrease in population pressure in Europe relative to China. In an alternative and complementary model, Gennaioli & Voth (2015) test the effects of war on state-building capacity in a two-player setting with European data from 1500 to 1800. They find that when the importance of money in war is low because technology is relatively simple, both contenders are similarly likely to win the war, so the richest ruler risks losing fiscal revenues and the poorer ruler consequently get a higher incentive to attack. In this case, which represents the Ottoman and Ming empires under frequent nomad raids or the European setting before 1650, warfare is frequent but the incentive to raise fiscal capacity is low. In scenarios where money is important for military success due to more advanced technology as in Europe after 1650, the odds of winning a war are stacked in favour of richer states that can afford state-building investments, whereas internally divided states find it costly to centralise and fall behind. Consistent with the European experience, a “race to the top” emerges, with all powers building state capacity as they compete more and more in fiscal and military terms.

Thirdly, political fragmentation in Europe might have drove European merchants to be more aggressive in their trade strategy and to compete with local elites for political powers; whereas elites in China and the Middle East had few institutionalised claims on the state and developed trade policies and institutions that maintained the existing social order. Cox (2017) argues that political fragmentation and representative institutions (self-governing cities and national parliaments) jointly determine the tariffs and regulations rulers could impose on trade; this in turn influenced inter-connected city growth. To support this hypothesis, Cox (2017) examined data sets on cities and their population from 800 to 1800 in the entirety of Eurasia to compute a distance weighted measure of inter-city growth correlations (which he claims correlates with traditional measures of market integration based on co-movements in prices, but also picks up deeper forms of economic integration, such as cooperation in specialised production chains). Cox (2017) then shows that inter-city growth correlations were negative and insignificant before 1200 and thereafter became consistently positive and significant only in Western Europe, the most politically fragmented and parliamentary region of all Eurasia. Within Western Europe, inter-city growth correlations were particularly strong in the most politically fragmented and fiscally strong parliamentary areas.31 In a similar line of thought, Blaydes & Chaney (2013) argue that the forms of executive constraint that emerged under feudalism in Western Europe were associated with an increase in relative political stability proxied by rulers’ lengths of reigns. Blaydes & Chaney (2013) found that reigns in Western Europe statistically diverged compared to the Islamic world around 1100. To explain this pattern, Blaydes & Chaney (2013) concentrated on the fact that feudal institutions in Europe served as the basis for military recruitment by monarchs, whereas Muslim sultans relied on mamlukism—the use of military slaves imported from non-Muslim

31 Cox (2017) notes that if his hypothesis is right, Europe began to diverge from the rest of Eurasia in terms of economic liberty as soon as the wave of communes and parliaments washed over the continent around 1200, but it is only c. 1300 that greater economic liberty generated a detectable advantage in per capita income and around 1800 that this advantage began to widen dramatically. He adds that several reasons could have hampered a more rapid Great Divergence, namely 1. endemic warfare; 2. monarchs’ efforts to suppress representative institutions, as soon as they could; 3. oligarchs efforts to use representative institutions to suppress rather than promote competition; and 4. the interfering features of markets themselves in a time before either private or social insurance was readily available.
lands. The authors then hypothesised that dependence on mamluk armies limited the bargaining strength of local elites with regards to the sultan, hindering the development of a productively adversarial relationship between sultanates rulers and local elites in the Middle East. Therefore, according to Blaydes & Chaney (2013), European rulers were already constrained compared to their Muslim counterparts by the time of the New World discovery and the associated “Colombian Exchange”.

3.3.3. The Atlantic slave trade and European colonialism

Both Chinese and Europeans rulers had a similar interest in maritime expansion before the fifteenth century. However, the Ming dynasty abandoned oceanic expeditions after a major political overthrow in 1424, while some European kings became increasingly interested in it (Morris, 2010, p. 413). In addition to these differences in motivations for maritime exploration, Morris (2010, p. 499) highlights that Europe was fortunate to have a decisive geographic advantage in reaching the New World since crossing the Atlantic was far more realistic with the existing nautical technologies compared to crossing the Pacific Ocean. Therefore, for this author, it is equally by accident than design that western Europeans and not Chinese created new kinds of oceanic empires. In the fifteenth and sixteenth centuries, Portugal and Spain pioneered European imperialism through the establishment of overseas colonies in South America, the Caribbean, and diverse coastal locations of Africa and India. In a second phase, France, the Netherlands, and Britain took the lead in colonising North America, the Caribbean, and different places in South and South-East Asia.

European colonialism and the associated slave trades are subjects that have received some of the most attention from those seeking to explain the institutional divergence of nations. Engerman & Sokoloff (1997) propose that, because of differences in the local climate and crop suitability, the different quality of institutions set up in various European colonies in the fifteenth century may have had a persistent effect on the level of development of countries once they achieved independence. Based on this idea, Acemoglu et al. (2001) report that in colonies without large European diasporas, such as Africa, Central America, the Caribbean, and South Asia, the objective was to oppress the native population and facilitate the extraction of resources in the short run. In colonies where Europeans settled in large numbers, such as the United States, Canada, Australia, and New Zealand, the institutions were being developed for their future benefits, and hence were inclusive. Acemoglu et al. (2001) further show that these different colonisation strategies were in part determined by the mortality rates of settlers. Furthermore, Acemoglu et al. (2002) report that Europeans were more likely to introduce extractive institutions in areas originally more densely populated by natives. Indeed, it was more profitable for them to exploit the indigenous population, either by having them work in plantations and mines, or by maintaining the existing system and collecting taxes and tributes. Finally, Acemoglu et al. (2001) argue that these early institutional differences have had long-lasting effects on present income per capita distribution. They find a significant positive correlation between the quality of early institutions and that of modern institutions, and a significant positive correlation between the quality of modern institutions and income per capita (when controlling for latitude, climate, current disease environment, religion, natural resources, soil quality, ethnolinguistic fragmentation, and present racial composition).

In India and Southeast Asia, European nations made more systematic use of commercial company-state to defend (in the figurative and literal meaning of the term) their mercantilist and imperialist interests. These include the Portuguese Estado da Índia, the French East India
Company, the Dutch East India Company, and the British East India Company (Findlay & O’Rourke, 2009). Although neither Parthasarathi (2011) nor Roy (2013) claim that pre-colonial India would have ever been on the verge of industrialising on its own, they disagree on the impact of English colonialism. For Parthasarathi (2011), pre-colonial India was relatively prosperous (even if it faced some ecological pressures and institutional barriers) so the increasingly tight hold of England clearly annihilated any chance of an early take-off. On the contrary, in Roy (2013) perspective, the Mughal empire was already weakening, probably with a slight decrease in GDP per capita, so the increasing interference of the English East India Company did not fundamentally hampered the Indian trajectory, and actually for this author the British rule might have reinvigorated industrialisation in India.

In the case of Africa, Nunn (2008) empirically investigates the long-term impacts of the successive slave trade episodes (through the Red Sea, Sahara, Indian Ocean, and Atlantic Ocean) from 1400 to 1900 on economic development in 2000. The author finds that the parts of Africa from which the largest number of slaves were taken are the poorest today. To control for the self-selection into the slave trades of societies with the worst performing institutions, Nunn (2008) examined the descriptive and quantitative evidence on the nature of selection during the four slave trades. He finds that it was actually the more developed and more densely populated societies of the time that supplied the largest numbers of slaves. Nunn & Wantchekon (2011) examine whether distrust could be the causal mechanism through which the slave trades have affected today’s economic development uncovered by Nunn (2008). The authors found a negative relationship between an individual’s reported trust in others (either neighbours, relatives, local governments, co-ethnics, and those from other ethnicities) in 2005 and the number of slaves taken from the individual’s ethnic group during the slave trades. Furthermore, Nunn & Wantchekon (2011) show that the lower levels of trust arose through two channels, namely 1. a deterioration of domestic institutions that enforce trustworthy behaviour and 2. an increase in the prevalence of cultural norms of distrust. The authors undertook several tests to distinguish between these two channels and the results suggest that the magnitude of the second determinant is about twice as important as the former. Finally, Michalopoulos & Papaioannou (2016) confirm the conventional belief that the Scramble for Africa (beginning with the Berlin Conference in 1884–1885 and ending in the early twentieth century), which artificially divided up the continent and partitioned ethnic groups, had detrimental impacts on long-run economic development through an increasing prevalence, severity, and duration of political conflicts.

Another consequence of the Atlantic slave trade concerns the underlying institutional causes of the British success in leading the Industrial Revolution. Acemoglu et al. (2005) argue that the growth of the Atlantic trade associated with European colonialism that started in the sixteenth century strengthened merchant groups by constraining the power of the European monarchies (until the overthrow of James II during the English Glorious Revolution). To support their hypothesis, Acemoglu et al. (2005) first show that the rise of Western Europe after the sixteenth century was driven by the economic growth of countries (and port cities) heavily involved in the Atlantic trade. Furthermore, they show empirically that among countries with access to the Atlantic trade, trade generated improved institution (measured using an index of a country’s constraints on its executive) only for those that had non-absolutist institutions initially (i.e., in the fifteenth and sixteenth centuries). If the monarchy was too strong initially, as in Spain and Portugal, it simply monopolised the trade, limited the benefits of the commercial class, and therefore limited institutional change. But if the monarchy was non-absolutist, as in England and the Netherlands, changes in relative bargaining powers helped merchants obtain changes in institutions to protect property rights.
(on land and capital), which paved the way for further innovations in inclusive economic institutions. As a result, English and Dutch merchant nations invested and traded more, but mostly England developed an institutional environment with its parliament that proved to be decisive to its industrial onset. Indeed, Mokyr (2011, pp. 486–487) speaks of the English Parliament as a meta-institution whose inherent flexibility was and remains crucial to sustaining economic growth by continuously adapting to the changing environment.

Furthermore, through the extensive use of slaves, the Atlantic trade allowed the extraction of natural resources and exotic products from the New World (sugar, tea, tobacco, coffee, fur, cotton, wood, and land-fertilising guano) which flooded Western European markets. Britain’s level of dependence on imports from the New World to carry out its industrial revolution is an old debate. Following Wrigley (1962), Pomeranz (2000, pp. 264–297) calculates the ghost acreages that would have been needed to feed and heat the British population of the nineteenth century if coal and natural resources from the American colonies (especially wood and sugar) had not been available. These calculations implicate that without the enormous consumption of coal to replace wood fuel, and the timber and calorie imports of the New World, Britain and other countries of Europe would have eventually faced an ecological bottleneck in the nineteenth century, which would have closed the industrial window and probably led Britain towards a Malthusian trap. As argued by Crosby (2004, pp. 2–7), the possibility of Britain to falling into a Malthusian trap in the nineteenth century was also considerably reduced by significant outward migration of its inhabitants toward North America. Theodoridis et al. (2018) estimate that if in 1832 Britain was indeed ghost-acreages reliant because land embodied in its imports was almost twice the land embodied in its exports, it was because of potash imports and not wood and sugar imports as emphasised by Pomeranz (2000).

Moreover, once the coal, iron, and steel industry increased sufficiently, the situation reversed and Britain became a net land-embodied exporter until the onset of the First World War. In any case, from 1500 onward, the expansion of European markets associated with the Atlantic slave trade, and the institutional changes that came along with it, has been important to lead several Western European countries towards an Industrious Revolution characterised by new forms of domestic consumption and an increasing prevalence of wage labor (i.e., labor offered on a specialised and relatively well-remunerated market instead of household labor) that together created a more natural transition to factory-based production (de Vries, 1994). Hence, in addition to having domestic institutions whose higher flexibility was revealed by the Black Death (see Pamuk, 2007, pp. 309–311), the imperialist and mercantilist momentum led by Britain and the Netherlands also contributed to the Little Divergence in real wages witnessed between the sixteenth and eighteenth centuries within Europe.

---

32 See Clark et al. (2014) for a review of the different arguments. In this paper, Clark et al. (2014) also use a computable general equilibrium model to run counter-factual simulations. The results support the idea that international trade only had a small effect on British welfare in the 1760s, but a very large impact in the 1850s. Consistent with Pomeranz (2000), the reason for Clark et al.’s (2014) results is that circa 1850, British population growth meant that the island depended on foreign agriculture for both food and raw materials, implying that it needed to export a growing amount of manufactures to pay for these imports. Also of interest in this debate is the argument made by Bateman (2012, p. 177) that the market alone is not a panacea to international trade, but in fact requires state interventionism. In contrast to their strong support for free trade, the economies experiencing the fastest growth in the nineteenth and early twentieth centuries (i.e., economies that are now among the richest of the world) were those using significant tariffs.

33 As noted by Vries (2016, pp. 26–30), the concept of Industrious Revolution has a quite different meaning for several authors, in particular for the inventor of the term, Hayami (2015), to whom it corresponds to the emergence of a labor-intensive agriculture compatible with the rise of a market in Japan.
3.3.4. Uneven coal endowment and the steam engine
A major implication of the intra-European Little Divergence is that incentives for labor-saving technologies increased from the sixteenth to the eighteenth centuries in Britain and the Netherlands compared to other early-modern European nations, while in China, Japan, or India it was nonexistent because labor remained relatively cheap (Allen, 2009, 2011, Allen et al., 2011). Moreover, because proto-industry relied heavily on wood fuel, wood scarcity (leading to price increases) was a frequent occurrence in most of Western Europe, especially in Great Britain (Pomeranz, 2000, pp. 220–223). As Goldstone (2002, p. 361) rightly states, the ultimate bottleneck in pre-industrial economies lays not in land or other raw materials but in energy. At these times of consequent incentives for both labor-saving and wood fuel-saving technologies in Western Europe, a natural accident had a decisive role. This natural accident consisted of the generous endowment of Western European countries, and here again most notably Britain, with its large and relatively accessible deposits of coal.34 But substituting wood and charcoal for coal for heat generation was well known for centuries and did not lead to an industrial revolution. What mattered, and constituted a major breakthrough, was being able to turn the heat from coal combustion into mechanical energy with the steam engine. Allen (2009, 2011) comprehensively argues that the British Industrial Revolution originated in the willingness and ability of its people to 1. tap their favourable coal endowment thanks to financial incentives represented in relative prices (of labor, capital, wood, and fossil fuels) and 2. apply knowledge from science (as already highlighted in subsection 3.2.4) to convert coal into mechanical energy, and in doing so direct and foster sustained technical change during the Industrial Revolution. Exploiting geographical variation in city and coalfield locations, alongside temporal variation in the availability of coal-powered technologies, Fernihough & O’Rourke (2014) support and quantify Allen’s (2009, 2011) theory. Precisely, Fernihough & O’Rourke (2014) estimate that coalfields’ location and the availability of steam engines explain around 60% of the growth in European city populations from 1750 to 1900.

In a similar vein, and after comparing the role of energy in Europe to other parts of the world over the past five centuries, Kander et al. (2013, p. 366) conclude that it is hard to imagine modern economic growth occurring without the adoption of fossil fuels, beginning with coal. They further emphasised that they view the transition to fossil fuels both as a necessary condition, and an enabling factor leading to modern growth (Ibid., italic emphasis present in original). Kander et al. (2013, pp. 367–368) then assert that the high complementary between coal, the steam engine, and the iron industry was crucial to deliver unprecedented amounts of mechanical energy that structurally reshaped the industrialising British society.35 They argued that the steam engine was one of the most important innovations in the history of mankind, because for the first time it was possible to reliably convert heat to motion in a controlled form, equipping people with inanimate “energy slaves” (machines). Steam engines

34 In China, coal deposits were distant from the major manufacturing regions of the lower Yangtze and the south (Pomeranz, 2000, p. 65).
35 It is because they completely miss the importance of the synergy between coal, the steam engine, and the iron industry that Clark & Jacks (2007), relying solely on the contribution of coal mining rents to national income, can claim that the possession of coal reserves made a negligible contribution to the success of the British Industrial Revolution.
enabled a large concentration of energy in time and space; they were more powerful than previous sources of kinetic energy and much easier to control. Steam engines saved labor and initiated a capital-deepening growth path where one worker could be in command of ever greater amounts of power. As Berg (1994, p. 207) notes, it was not the spinning machinery itself (in operation since 1770) that made England a leader in textile production, but rather the application of steam power to spinning, water and surface transport, brick-making, grain-threshing, iron-making, shovelling, construction, and all sorts of manufacturing processes that transformed the British economy. Hence, the entire process of capital-deepening growth was almost wholly reliant on fossil fuels and eventually led to both increased incomes and a dynamic that has continued to raise incomes (Kander et al., 2013, p. 368). In line with this reasoning, Goldstone (2002, p. 362) rightly points out that even if China certainly had all the raw materials and all the needed technology, it did not accomplish an efflorescence based on adapting water power to the mechanical manufacture of cotton textiles. The explanation for the absence of such a water power-based efflorescence is still controversial, but in any case, Goldstone (2002, p. 363) notes that even if China had made the step of deploying water-powered spinning, that by itself would not have launched an industrial revolution. To do so would have required a systematisation of knowledge regarding air pressure, power, and its application to constructing heat engines. Unlike water-powered spinning, there is no evidence of such an independent trajectory in China.

To quantify the importance of coal as a source of both heat and mechanical energy in the transition from limited to sustained economic growth, Malanima (2016, pp. 96–99) improves and extends Pomeranz’s (2000) estimates of land- and labor-savings resulting from coal use in England and Wales during the period 1560–1913. The results showed two distinct historical phases. During the first one, which lasted from the end of the sixteenth century until about 1830, the use of coal was mainly land-saving. It is only during the second phase (from 1830 to 1900) that coal was both land and labor-saving. Covering both phases from 1800 to 1900, the land-related (resp. labor-related) savings grew from 1 to 14 times the extent of the entire country, that is 15 million hectares (resp. from 1 million to almost 290 million workers in 1900 when the English population was 32 million and the labor force 13–14 million). These estimates strongly support Wrigley’s (2016, pp. 2–4) claim that the energy required to produce iron and steel on a large scale or to construct and operate a railway system implied that it was unreasonable to expect that it could be secured from the annual flow of energy derived from plant photosynthesis. As a corollary, for Wrigley (2016), an industrial revolution could not have been accomplished as long as mechanical energy continued to be provided principally by human and animal muscle.

36 It is often argued that early steam engines were extremely inefficient and that the end-use cost of the mechanical energy produced by a steam engine was not lower than that produced by a windmill, waterwheel, or worker. This so-called paradox is solved when one considers that, unlike the prime movers that preceded it, a steam engine is both mobile and continuously operational. A calculation of the production cost for the same “controllable energy unit” would clearly show the net advantage of coal combined with a steam engine compared to all competing prime movers of the same period.

37 As noticed by Malanima (2016, p. 96), usual social savings calculations based on relative costs of old and alternative technologies appear quite impossible here because it would require to compute counter-factual wood prices and labor wages in a theoretical British economy where coal would have been absent.
4. Reflections on persistence and reversal of fortune

4.1. Persistence of fortune: technology and culture in the very long run

Several studies support the idea of the persistence in economic development over significant periods of time. The best study supporting long-term persistence of discrepancies in technological and economic development is provided by Comin et al. (2010). First, the authors assembled a new dataset on the history of technology from 1000 BCE to the “Age of Discovery” circa 1500 CE. Then, Comin et al. (2010) showed that technological differences between the predecessors to today’s nations identified as long ago as 1000 BCE were still observable and even widened in the years 0 and 1500 CE. Technology differences in 1000 BCE and 0 CE robustly predict technological gaps in 1500 CE. Moreover, technology in 1000 BC and 0 CE is also sometimes significant as a predictor of income and technology today (2002), but these associations are not robust. However, Comin et al. (2010) showed that the 1500 CE measure of technological differences is a statistically significant predictor of the pattern of per capita incomes and technology adoption across nations observed nowadays (2002). The authors improved the results by using migration data from Putterman & Weil (2010) to have a population-weighted average places of origin of the technology. As a possible causal mechanism underlying their empirical results, Comin et al. (2010) found evidence that technology adoption dynamics are driven by increasing returns (the cost of discovering and adopting new technology falls with the stock of previous technologies).

Less directly related to technology and economic development, but still relevant for the Great Divergence debate, Alesina et al. (2013) tested the hypothesis that traditional agricultural practices influenced the historic gender division of labor and the evolution of gender norms. They find that the descendants of societies that traditionally practised plough agriculture today have less equal gender norms, measured using reported gender role attitudes and female participation in the workplace, politics, and entrepreneurial activities. Using here again the international migration data of Putterman & Weil (2010), Alesina et al. (2013) showed that among the children of immigrants born and raised in a European country or the United States, those with a heritage of traditional plough use exhibited less equal beliefs about gender roles today. This shows how embedded this long-term cultural persistence can be.

4.2. Reversal of fortune: critical junctures and turning points

Contrary to studies showing long-term persistence, several articles emphasise that because of critical junctures brought by historical contingency, certain sets of events can radically shift pre-existing development paths. The advent of the Black Death is the typical example of such a major turning point at the continental scale. As we saw in Section 3.3.1, the Black Death and its subsequent redundant episodes had profound effects on European countries, revealing the differences they had in terms of domestic institutional flexibility. Acemoglu et al. (2002) report a reversal of fortune among colonised territories of the Americas: the most prosperous native American locations in 1500 locations are the most underdeveloped today after being colonised by European countries.38 Also, recall the reversal of fortune documented by Nunn

38 Williamson (2011) judiciously points out that globalisation since 1500 and Western industrialisation in the nineteenth century meant local deindustrialisation for their colonies (that were to become the “Third World”) because of comparative advantages in the exploitation of natural resources. However, a focus on natural resource extraction has not been an ineluctable barrier to the modernisation of nations such as Australia, Canada, New
(2008) who showed that among African countries, those that were the most developed prior to the slave trade (measured by population density in 1400) had the largest number of slaves taken and have the lowest economic output today.

On a longer time span, Borcan et al. (2018) analysed how state development since 3500 BCE has interacted with economic development until the present day. The authors provide a complete state history index for 159 contemporary countries running from 3500 BCE to 2000 CE with a 50-year period. Borcan et al. (2018) showed that the relationship between state history and current income per capita across countries is hump-shaped (i.e., has the shape of an inverted U) rather than linear, and that this is due to the inclusion of state experience before the common era. Thus, in addition to young inexperienced states (e.g., Mauritania, Zimbabwe, Haiti, and Estonia), very old states (e.g., Iraq, Turkey, Egypt, and China) also incurred economic disadvantages relative to intermediate states with less than 2000 years of state experience (e.g., Britain, Italy, Denmark, and Japan). The authors also showed that the non-linear impact of state history is already visible with respect to economic development indicators (population density and urbanisation) and technology adoption in 1500 CE, but it clearly reveals itself in the economic performances of 2000 CE.\footnote{To explain their result, Borcan et al. (2018) argue in a stylised theoretical framework that the earliest states developed the fiscal capacity and coordination needed to achieve increases in productivity, but ultimately limited that productivity due to over-centralisation. Although earlier states became stagnant, younger states were able to learn from them and surpass their productivity before they reached stagnation themselves. By contrast, very young states early in the process of building fiscal and institutional capacity are at a relative disadvantage. Thus, along with young states, a very long state experience also comes with economic disadvantages relative to countries with intermediate state experience.}

In addition, the relationship for current outcomes is further strengthened when adjusting the index for the ancestral lines of post-1500 migrant populations provided by Putterman & Weil (2010).

Furthermore, reversal of fortunes can be witnessed within a development mechanism that would appear to be the quintessential example of long-term persistence. In Section 3.1.3 we show how the timings of the transitions from foraging to farming in the Neolithic period have had long-term development impacts observable even as late as 1500. However, Olsson & Paik (2013) show that the parts of Europe that first adopted agriculture, and were arguably more economically developed during the Neolithic period, are less developed today (they also find evidence of a similar reversal within sub-Saharan Africa and East Asia). To explain this reversal of fortune, Olsson & Paik (2016) claim and empirically demonstrate that the advent of farming in Mesopotamia was characterised by collectivist values, which after a while would sooner or later triggered the out-migration towards more and more peripheral areas of Western Europe of the more individualistic farmers that would eventually generate the emergence of institutions more conducive to economic growth.

4.3. So, persistence or reversal?

There seems to be a contradiction between studies highlighting persistence of development discrepancies over millennia, and those pointing out to abrupt reversal of fortunes. So, which is correct? It turns out that both are, and reconciling their findings requires noting that they generally differ in the samples being examined: the persistence studies tend to examine all countries globally, while the reversal studies only include sub-global or sub-continental samples (Nunn, 2014, p. 387).
Chanda et al. (2014) brought some compelling evidence to this explanation. Using the 1500–2000 world migration data of Putterman & Weil (2010), Chanda et al. (2014) confirmed the reversal of fortune for colonised countries with territories as the unit of observation, but once they use people and their descendants as the unit of observation, they find instead a persistence of fortune. The authors show that their results are robust to the choice of end year (1960 versus 2000), countries sub-samples, and four alternative measures of early development in 1500 (namely, population density, time since transition to agriculture, history of state-level polities, and technology level). Furthermore, Chanda et al. (2014) found supportive empirical evidence to Glaeser et al.’s (2004) ideas that 1. Europeans who settled in the New World may have brought with them not so much their institutions, but themselves, that is, their culture and human capital and that 2. human capital has been a factor at least as important as (if not more fundamental than) institutions in determining long-term comparative development.40

Similarly, although Olsson & Paik (2013) found a reversal of fortune within Europe from the Neolithic period until the present day (parts of Europe that adopted agriculture earlier, and were arguably more economically developed during the Neolithic period, are less developed today), with a global sample they find persistence (parts of the world that adopted agriculture earlier are more developed today) similarly to previous studies (namely, Ashraf & Galor, 2011, Olsson & Hibbs, 2005, Puttermann, 2008, Putterman & Weil, 2010).

Finally, the study of Kelly (2019) must be mentioned to finish this discussion on the persistence and reversal of fortune. This work has not yet been published in a peer-reviewed journal, but it points to potentially serious flaws in long-term persistence studies. Precisely, Kelly (2019) analyses 27 persistence studies (among which 14 are cited in the current paper) and find that in most cases if the main explanatory variable is replaced with spatial noise the fit of the regression commonly improves, and if the dependent variable is replaced with spatial noise, the persistence variable can still explain it at high significance levels. These results suggest that the significance levels of a great deal of persistence studies, and of spatial regressions more generally, might be overestimated.

5. Conclusion: path-dependence in a contingent world

Establishing historical causality is usually complex and the economic sphere is no exception to this rule (Ville, 2015, p. 85). In Section 2, I highlighted the uniqueness of the Great Divergence phenomenon with respect to past efflorescences and I also examined the most recent GDP per capita estimates to determine the respective timings of the Little and Great Divergence. The beginning of the Little Divergence between Western Europe on the one hand, and South and Continental Europe on the other, is visible in GDP per capita and real wages around 1350–1400, whereas the occurrence of the Great Divergence should be dated from 1700 to 1750.

Past and ongoing debates reviewed in Section 3 have allowed a detailed examination of the occurrence and persistence of the Great Divergence. Several biogeographical factors (climate conditions, continental sizes, orientation of major continental axes, length of coastline

---

40 To support their point of view, Glaeser et al. (2004) argue that cases of developing countries escaping poverty through sound policies enacted by dictators exist and that the positive economic outcomes of such a transition then improve political institutions. Glaeser et al. (2004) recognise that their view is clearly in line with the “Modernisation theory” developed by Lipset (1960). This approach suggests that economic growth and the processes that go along with it, such as expanding education, urbanisation, or the development of a middle class, determine institutional change, and not the other way around. Lipset’s (1960) hypothesis received substantial empirical support from Barro (1999) and Przeworski et al. (2000), and it was restated by Friedman (2005).
compared to mainland area) have had an undeniably deep influence on the timing of the agricultural revolution of the Neolithic, which then had long-lasting effects on the differential development of world regions. The initial advantage of Eurasia loosened over time, but circa 1000–1350 it was already more probable to bet on a future economic take-off toward sustained growth in that part of the world rather than elsewhere. Culture and its social implementation through institutions have then had a critical role in determining the different scientific and technological trajectories of world regions. Despite major technological discoveries, the successive Islamic and Chinese empires gradually developed monolithic and relatively rigid societies in contrast to the more pluralistic and flexible institutions of European countries. In particular, from the sixteenth to the eighteenth century, Europe’s decentering and repudiation of its classical religious and philosophical traditions in the study of nature has had no counterpart in any other major civilisation prior to the nineteenth or even twentieth century (Goldstone, 2012). Hence, European intellectuals had more incentives to discover the natural laws of the world, and the emergence of modern science in the technologically backward Europe of pre-industrial times seems to have been influential in determining the sub-continental location of the Industrial Revolution.

Hence, from 1350 to 1750, a future economic take-off towards modern growth was increasingly likely in Western Europe. Over this period, England was certainly not an uncontested candidate as Italy and then the Netherlands were the leading regions of Europe. Cultural–institutional deep determinants are effective in a certain context so that their long-lasting effects do not imply an absolute determinism. Indeed, as put by Mokyr (2011, p. 486), ideology is an integral part of economic change, but just as there is no fixed set of “good” institutions that are suitable for the economy under all circumstances, there is no “right” ideology that works in all circumstances toward economic progress. Therefore, it is also a succession of accidents and contingent historical events (the Black Death, the intra-European fragmentation, the abandonment of oceanic expeditions by the Ming dynasty, the Atlantic and Pacific relative sizes, the slave trades, China’s need for silver, and the location of coal deposits) that worked in conjunction to explain 1. the earliness of Britain to progressively reach a modern regime of economic growth that would set in motion the Great Divergence, 2. the interventionism of strong government in Western Europe to promote mass literacy and protect domestic markets through tariffs, and 3. the establishment of extractive institutions in European colonies that would ensure (with other determinants) the persistence of the Great Divergence.

In summary, it appears that the Great Divergence was never an inevitability but became an increasingly likely prospect as time progressed. Moreover, biogeography, culture-institutions, and contingency-conjuncture are not contradictory hypotheses. Rather, there is a clear pattern of change over time of the relative importance of these three categories of determinants. The relative importance of each of this category varied successively over time, but none of these class of factors uniquely determined the course of the Great Divergence. Therefore, in addition to studies that deal with the long-lasting effect of a (set of) factor(s) in the context of a particular historical event, research should concentrate on the elaboration of a unified framework able to account for the relationships of the different determinants of economic development. Such a unified theory is needed to understand which causal link or latent variable acted between biogeographical, cultural–institutional, and contingent–conjunctural determinants to explain their successive relative importance over time.
Acknowledgments

This work benefited from the support of the Chair Energy & Prosperity. I thank Pierre-Olivier Pineau, Adrien Nguyen-Huu, and Michel Lepetit for their helpful comments on earlier versions of this article. I am also grateful to three anonymous reviewers and Editor-in-Chief Nikolaus Wolf for their fruitful comments and suggestions. Many thanks to David Eggleton for correcting the spelling and grammar of this text. All remaining errors are mine.

Conflict of interest statement: The author declares that he has no competing interests.

References


McCLOSKEY, D.N. (2008). You know, Ernest, the rich are different from you and me: a comment on Clark’s “A Farewell to Alms”. European Review of Economic History 12, pp. 138–148.


