TRADE AND THE PATTERN OF EUROPEAN IMPERIALISM, 1492-2000

ROBERTO BONFATTI
Trade and the Pattern of European Imperialism, 1492-2000

Roberto Bonfatti
Oxford University

April 30, 2012

Abstract

I construct a trade model of empire, and use it to interpret some of the key patterns in the history of European imperialism. I begin from the observation that trade was a key source of wealth for the colonies, and trade restrictions a key tool of extraction for colonial powers. But the value of this tool must be seen in relation to the value of colonial trade, and to the extent of international competition for it. The model interprets the colonial empires that emerged in the 16th-18th century as a set of political institutions designed to appropriate the value of colonial trade to the mother country, at a time in which colonial trade was both valuable and highly competed for. It explains the fluctuations in the fortunes of empire in the 19th and early 20th century with the rise of a clear industrial leader, Britain, and her subsequent decline. Finally, it attributes the fall of colonial empires to a secular fall in the importance of colonial trade, relative to trade between the industrial countries. I provide detailed historical evidence in support of these predictions. The model also has predictions for the impact of empire-building on trade relations between the imperial powers. These are consistent with the apparent inverse relation between European imperial expansion and globalization.

JEL Codes: F1, F5, N4.

Keywords: Imperialism, Preferential Trade Agreements, Intra-Industry Trade, Hegemonic Stability.

1 Introduction

The rise and fall of European empires represents one of the key facts in modern history. European imperialism shaped the world almost five centuries, leaving a long-lasting economic legacy. The

*I would like to thank Tim Besley, John Darwin, Avinash Dixit, Steven Pincus, Steve Redding, John Romalis, Kevin O’Rourke, Tim Schmidt-Eisenlohr, Tony Venables, Jeff Williamson and Nuala Zahedieh for useful comments, as well as participants at the Oxford Trade Workshop (Feb 2011), the Oxford Economic and Social History seminar (May 2011), the European Trade Study Group (Sep 2011) and the Leuven Workshop on the History of Globalization (October 2011). All remaining errors are my own.
latter has been the subject of growing debate in economics, sparked by the work of Acemoglu, Johnson and Robinson (2001).

To better understand the impact of imperialism on past and present economic conditions, it is important to understand the forces that led to its rise, evolution, and decline over time. Why did Europeans build centralized empires from the 16th to the early 20th century? Why did they do so in some periods more than in others? And finally, why did empires fall in the later 20th century? This paper studies one central aspect of this - the influence of trade. Colonial trade emerged as an important source of wealth from the 16th century, and a key benefit of empire was that it allowed the mother country to divert that wealth from colonies and from competing European powers. But the value of empire must then be seen in relation to the value of colonial trade, and to the intensity of competition for it. I set up a model to make this idea precise, and use my results to show that the pattern of imperialism may be explained with changes in the value of empire, which in turn were determined by changes in the pattern of trade. In particular, the rise and secular decline of European imperialism may be explained with the rise and decline of colonial trade, relative to trade between European countries. And fluctuations in the pace of imperial expansion may be explained with fluctuations in the intensity of European competition for colonial trade. The model also provides an explanation for the fact that imperial expansion was associated with periods of restricted trade between European countries, while imperial relaxation or contraction was associated with periods of globalization.

Any theory attempting to provide a consistent explanation for the rise and decline of European imperialism faces the challenge of a very articulated pattern over the period 1492-2000. Figure 1 illustrates this pattern by plotting the estimated areas of overseas territories directly administered by Europeans\(^1\) in this period. To the discovery of the Americas (1492) and the sea route to the Indies (1498)\(^2\), followed a first long period of expansion (Phase I: 1492-1770), during which Portugal, Spain, and later the Netherlands, England and France constructed colonial empires both in the West and in the East. After a rapid contraction - mainly driven by colonial rebellions

---

\(^1\)Plus the US and Japan. Throughout the paper, I interchangeably refer to “European” and “Western” imperialism.

\(^2\)Although the early African ventures by the Portuguese (such as the conquest of Ceuta in 1415) are considered by historians as the real beginning of European imperialism, these were of little economic importance relative to the two major breakthroughs of the late 15th century.
in the Americas - there was a period of about 50 years (II: 1830-1880) in which Britain was left as the only significant colonial power. One striking fact about this period was that despite being left a clear military leader, Britain did not enlarge much her directly-administered empire, and she agreed to concede responsible government (a substantial form of autonomy) to most of her colonies of settlement.\(^3\) The late 19th century marked the beginning of a new phase of territorial expansion (III: 1880-1938), during which old and new European empires (plus Japan and the US) expanded to dimensions never seen before. Finally, all colonial empires collapsed in the several decades after the end of WWII (IV: 1938-2000). While various reasons for this collapse have been put forward, one view holds that, with a few clear exceptions,\(^4\) decolonization was motivated by the fall in economic profitability of empire.\(^5\)

---

\(^3\)These were Canada, Australia and New Zealand.

\(^4\)These include Algeria and the Portuguese colonies.

\(^5\)There are two main competing explanations. The first is that the European metropoles lost their relative power, and were induced to give up their empires by US pressure. The second is that nationalists movements became too strong to be kept under control, therefore inducing decolonization. I discuss these alternative explanations in section 4.

---

Figure 1: **Areas of overseas territories directly administered by the imperial powers, 1492-2000 (M sq. km).**
Source: Etemad (2007, p. 131) and author’s calculations (all details are in Appendix C).
But what contemporary economic conditions were these large swings in European imperialism associated with? One fact that immediately catches the eye is that there seems to be a direct relation between the intensity of imperialism and the level of trade restrictions between the imperial powers. On one hand, imperial expansion was historically associated, or closely followed, by periods of highly restricted trade between the imperial powers. This is particularly evident in the latter part of Phase I (say 1650-1770), which is the period in which most of the Dutch, English and French expansion took place. This period is often called “the age of mercantilism”, because it was characterized by pervasive restrictions in the trade between European countries and their empires. It is also evident from the early years of Phase III, which marked the beginning of a process of increase in tariffs among European countries, that would culminate in the disastrous protectionism of the 1930s. On the other hand, imperial relaxation or contraction were associated with periods of globalization. This was the case in the central decades of Phase III (1840-1870), when, on British lead, a series of bilateral treaties were signed between European countries that created the basis for the first wave of globalization. And it was the case in Phase IV, which witnessed the creation of GATT/WTO, the multilateral institution that underpin the current wave of globalization.

I develop a trade model of imperialism, and use it to construct an argument explaining the pattern described above. In a nutshell, the argument is as follows. The great discoveries of the late 15th century opened up a new inter-continental trade for Europeans. This trade - mostly an exchange of raw materials/foodstuff for European manufactures and services - had the characteristics that Europe’s trade partners were relatively weak from a military point of view, and could therefore be incorporated into colonial empires. The latter were institutions that, among other things, allowed the mother country to capture the value of colonial trade, to the detriment of colonies and of European competitors. Colonial trade remained important for the European economy until the first half of the 20th century. Throughout Phases I, II and III, there was then a force that made imperial institutions potentially profitable. However this also depended on the intensity of European competition for colonial trade. Phases I and 

---

6 See, for example, Ormrod (2003).
7 There were important exceptions, such as the export of silver from Europe to Asia from the 16th to the early 19th century, and the import of Indian textiles into Europe from the late 17th to the early 19th century.
III were characterized by strong competition.\textsuperscript{8} This made imperialism highly profitable for multiple European countries, leading to frequent colonial confrontations or wars, and to frequent annexations. On the contrary, Phase II was associated with a clear dominance by Britain of colonial markets. This reduced the incentives for Britain to try to secure commercial privileges through imperial control, creating an equilibrium in which overwhelming British power was used to patrol the high seas, but not to massively annex new territories. With the rise of intra-manufacturing trade in the second half of the 20th century, colonial trade declined in importance relative to trade in manufactures between the industrial nations. This decreased the profitability of imperialism, contributing to explaining its fall (Phase IV). Finally, imperial expansions in Phases I and III created larger “domestic” markets for imperial manufactures, and ones in which consumers, being partly colonial, had a lesser weight in government objectives. This increased pressure for protection of domestic markets, and may have contributed to the rise of intra-European tariffs in the 18th century and in 1880-1938.

Turning to the mechanics of the model, one way to look at international trade in the 17th-early 20th century is that Europeans\textsuperscript{9} competed to export mostly homogeneous manufactures and services to the rest of the world, in exchange for raw materials and foodstuff. On the contrary, the second half of the 20th century witnessed a rise in the importance of differentiated manufactures, that are exchanged not only with the rest of the world, but also among European (advanced) economies.\textsuperscript{10} I thus proceed in two steps. First, I set up a simple 2-product, 3-country Heckscher-Ohlin model, in which two advanced countries compete to sell a homogenous good $y$ to a potential colony abundant in a homogenous good $x$. This model features no trade between the two advanced countries. Second, I extend this simple set up to allow for Krugman-style differentiated varieties of good $y$, generating (intra-industry) trade between the two advanced countries as well. The two advanced countries can choose to be imperially active, which gives them the option of annexing the potential colony at an exogenous cost $k^A$. If they are both

\textsuperscript{8}This was mostly between pre-industrial Holland, England and France in Phase I, between Britain and various industrial followers (Germany, France, Italy, the US, Japan) in Phase III.

\textsuperscript{9}This includes the US and Japan from the second half of the 19th century.

\textsuperscript{10}There was, to be sure, a lot of trade between European countries in the 17th-20th century as well. Intra-industry trade in manufacturing and services was very small, however, implying that extra-European markets were relatively important for competing European exporters.
active, however, a war ensues, which entails an exogenous cost $k^M$. I assume that the war is more likely to be won by the country that has the highest production of $y$, or that is more advanced. Following to a war, the winner has the option of annexing the potential colony, always at a cost $k^A$. I assume that the only implication from annexing the colony is that the imperial power can then control her trade policy. This entails two gains for the imperial power. First, she can restrict access to the colonial market to its competitor, thus obtaining better terms of trade. Second, she can prevent being herself excluded from the colonial market.

In the simple model with homogenous $y$, these gains are always rather high, since the colonial market is an important outlet for the two advanced countries. Whether or not this results in actual empire building, however, depends on the relative development of the two advanced countries, as producers of $y$. When the two are rather similar, they have both comparable military strength and comparable shares of the colonial markets, the latter implying that they benefit a lot from excluding their competitor. This creates an equilibrium in which both countries are imperially active, there is war, and the potential colony is always conquered (Phases I and III). On the contrary, when there is a clear leader, this country is both military powerful and has a high share of colonial market, which implies that she has little to gain from excluding her competitor. This creates an equilibrium in which the leader is alone in being imperially active, but does not annex the potential colony (Phase II).\footnote{Notice that the leader must be active in this equilibrium, or else the follower, which has a little share of the colonial market and would therefore gain a lot from excluding the leader, would step in and annexing the potential colony.} In the extended model with heterogeneous manufactures, the importance of the colonial market as an outlet for $y$ is decreasing in the degree of differentiation of alternative varieties. There then exists a degree of differentiation for which the gain from excluding competitors from colonial markets is low for all levels of competition, and it is more than compensated by the loss from restricting imperial consumption of other varieties of $y$. The only equilibrium is then one in which there are no colonial wars, and empires are not constructed (Phase IV).

Turning to the related literature, any economic theory of imperialism must first of all be seen in its relation with competing theories of imperialism, both economic and non-economic. Here I limit myself to summarizing what I believe are the three relative strength of my approach. First,
it builds on a mainstream historical view of European imperialism that puts the economic forces and trade at centre stage. Second, while it cannot by itself explain the experience of each specific colony, it does well in explaining some of the global patterns of imperialism. Finally, it provides one consistent story as to why imperialism arose, evolve over time in some specific ways, and eventually fell. This is in contrast with theories which, while providing a appealing explanations for a specific phase of imperial history, fail to explain the various other phases.

The paper is also related to several strands of literature on international trade, international relations, and history. I review these three fields in turn. Essentially, the colonial trade restrictions on which I focus amounted to preferential trade agreements between the colonies and the mother country (where the latter received most of the benefit). For this reason, the paper is strongly related to the theoretical literature on preferential trade agreements (e.g. Grossman and Helpman, 1995; Bagwell and Staiger, 1999a; and Ornelas, 2005). Next, the paper relates the fall of colonial empires after WWII to the contemporary liberalization of trade among the industrial countries. For this reason, it is connected to the literature on the economic origins of the WTO (e.g. Bagwell and Staiger, 1999b; Ossa, 2011). The paper is very related in spirit to a literature on trade and conflict, in the presence of insecure property rights (e.g., Skaperdas and Constantinos, 2001; Garfinkel et Al., 2011), but differs from these in that it specifically focuses on the establishment of imperial institutions. Finally, the paper is complementary to an earlier work by the same author (Bonfatti, 2008), investigating the economic incentives for colonies to rebel to their mother country.

Within the field of international relations, the paper is very related to an area known as “hegemonic stability theory”. This line of research - pioneered by the work of Charles Kindleberger (1973) - contains several theories\(^\text{12}\) about how the presence of an economic hegemon (Britain in the 19th century, the US in 20th century) stabilizes international relations. Furthermore, many of these theories explicitly predict that the hegemon is not interested in extending formal control over foreign territories, which is also one of the key results of my paper. As for the relevant historical literature, further details on this are provided in section 4. Here I limit myself to

\(^{12}\)Among the main hegemonic stability theories are realist theories (Keohane, 1984; Gilpin, 1975,1989; Krasner, 1978), world-systems theories (Wallerstein, 1974-2011) and long-cycle theories: Modelski (1989).
observing that the implications of my paper for the history of the British Empire are consistent with the findings of a few key pieces of work on this topic (e.g. Gallagher and Robinson, 1953; Cain and Hopkins, 2001).

The paper is organized as follows. In section 2, I setup the baseline model with homogeneous manufactures. I extend this to a case with heterogeneous manufactures in section 3. In sections 4 and 5, I use the model to interpret some of the long-term pattern of European imperialism. Finally, in Section 6, I conclude.

2 Baseline model

Consider a world made up of three countries, called B, F and C. Countries B and F are potential imperial powers (we shall think of them as Britain and France) while C is a potential colony. There are two tradable goods, x and y, which could be either final or intermediate goods. Preferences are described by Cobb-Douglas utility (or by linear utility defined over a final good \( z = x^{\frac{1}{2}}y^{\frac{1}{2}} \)):

\[
U = x^{\frac{1}{2}}y^{\frac{1}{2}}.
\] (1)

Goods exist in the form of endowments, and are distributed across countries as follows:

\[
\begin{align*}
x^B &= 1 \\
x^F &= 1 \\
x^C &= 1 \\
y^B &= b \\
y^F &= f \\
y^C &= 0
\end{align*}
\]

where \( b, f \in [1, 2] \). I assume that, within each country, individual endowments are small, so that no one can single-handedly affect prices. Clearly, the specific distribution of endowments that I have chosen makes both B and F abundant in y, relative to C. We can then anticipate that both countries will be exchanging y for x with C. To build intuition, I will from now refer to x and y as, respectively, “raw materials” and “manufactures”. However I will later adopt a more flexible approach, in particular referring to y as “services” as well.

The equilibrium of this simple model has a very attractive feature. Because of Cobb-Douglas
utility, relative demand for the two goods is always equalized to the inverse of the price ratio. It follows that in equilibrium, the inverse of the price ratio must equalize relative supply. Thus, suppose that free trade realizes among a subset \( M \) of countries. Normalizing the price of raw materials to one, the equilibrium price of manufactures within that set must be:

\[
p^M = \frac{x^M}{y^M}
\]

(2)

where \( x^M = \sum_{J \in M} x^J \) and \( y^M = \sum_{J \in M} y^J \) are the total supply of the two goods within the set. Notice that it is implied from equation 2 that the autarchy equilibrium of country \( J \) is simply \( p^J = \frac{1}{y^J} \).

### 2.1 World-wide free trade equilibrium

Suppose that free and costless trade takes place among the three countries (thus, \( M = W \equiv \{B,F,C\} \). The import pattern of each country is found to equal:

\[
m_x^J(W) = \frac{p^W y^J - 1}{2}
\]

\[
m_y^J(W) = \frac{1 - p^W y^J}{2p^W}.
\]

Equations 3 and 4 clearly indicate that, in the world-wide free trade, a country exports manufactures if their free trade price is higher than their autarchy price, it imports them otherwise. But because it is always \( \frac{3}{b+f} \geq \frac{1}{b}, \frac{1}{f} \) in our parameter space, \( B \) and \( F \) are always exporters of manufactures,\(^\text{13}\) which implies that \( C \) is always an exporter of raw materials. The following figure confirms this pattern, by plotting trade in raw materials in three alternative scenarios: \( b \) and \( f \) increasing at the same time, \( b \) increasing with \( f \) constant at \( f = 1 \), and \( f \) increasing with \( b \) constant at \( b = 2 \).

As figure 2 illustrates, the model describes a world in which the two potential imperial powers compete for importing raw materials from the potential colony, in exchange for the export of

\(^{13}\)Except for the case \( b = 2, f = 1 \) (or vice versa), in which \( B \) (\( F \)) exports manufactures while \( F \) (\( B \)) does not trade. This limiting case will play little role in the subsequent discussion.
manufactures or services. This is a reasonably good description of the new trade that the discoveries of the late 15th century opened up between Europe and both the East (the Indian Ocean) and the West (the Americas). In the 16th-18th, a group of Atlantic European countries (Portugal, Spain, and, later, the Netherlands, England and France) competed to import spices from the East and valuable commodities such as sugar, coffee or chocolate from the West. While in the East these were largely paid for in silver, in the West the settlements of the new territories provided substantial markets for the manufactures produced in each of the Atlantic European economies (and particularly in the Netherlands, England and France). Both in the East and in the West, the Atlantic European economies competed to export shipping and intermediation services to the New World, since most of the imported raw materials were actually eventually bound for the broader European market.

Another important feature of the parameters is that they allow for the full range of degrees of competition between $B$ and $F$. On one hand, competition is maximum when the two countries produce an equal amount of $y$ ($b = f$), or are at a similar level of “industrial development”. This is because they each control exactly half of $C$’s market in this case. On the other hand, competition is decreasing in the difference in industrial development between the two countries, in the sense that the higher is this measure, the more disproportionate is the share of the industrial leader in $C$’s market. To the limit, when the distance between $b$ and $f$ is maximum, the industrial leader controls the whole of $C$’s market, and the follower does not trade at all. Where the difference in industrialization to increase even further, the follower would become itself an exporter of raw materials, thus competing with $C$ for the leader’s market.
One may wonder why C’s exports of raw materials does not increase as B or F (or both) become more industrialized. After all, industrialization normally leads to a higher demand for raw materials, both produced domestically and imported. However notice that industrialization has always two opposite effects on the demand for \( x \): a positive one, through a higher income in the industrializing country; and a negative one, through a higher relative price of raw materials (Figure 3). These two effects exactly offset in this model, leading to a constant export flow of raw materials. Still, industrialization is associated with a higher volume of trade, since a constant export flow of raw materials corresponds to a larger import flows of manufactures as the price of these has decreased. Figure 4 illustrates that the share of trade in GDP is always non-decreasing for an industrializing country.

Figure 3: **Relative price of manufactures and industrialization.**

Figure 4: **Trade to GDP ratio and industrialization.**

\( M^J \) is country \( J \)’s imports plus exports, and both \( M^J \) and \( GDP^J \) are expressed in terms of the numeraire good.
2.2 Trade policy when all countries are independent

Suppose now that each country is governed by a welfare-maximizing government, which sets national trade policy. In this world with two goods and three countries, a complete trade policy vector would be defined as a pair or import tariffs (or, alternatively, export taxes) that each country raises on trade with each of the other two countries. I simplify things by substituting this continuous policy space with a discrete one, in which each tariff can either be zero or prohibitive (infinite). In other words, a government can only decide whether to trade freely with a certain country, or block (direct) trade with it altogether. While this assumption reduces drastically the realism of the model, it does allow me to illustrate the main points in a very simple way.

We begin by considering a world where \( C \) is an independent country, whose government is free to set its own trade policy. Because I will define “imperialism” as a transfer of trade policy-setting powers from the colony to the imperial power, the current scenario is the relevant counterfactual to the centrally-administered European empires that emerged in the 16th-18th century. It is the scenario that would have realized if, on one hand, the trading nations of Asia had not been progressively brought under the administrative control of the countries of Atlantic Europe. On the other hand, it would have realized if the European emigrants who settled the Americas had been granted a substantial degree of autonomy (or outright independence) vis-a-vis their mother country. It is, of course, the scenario that realized after decolonization, and in the few British settlers’ colonies that received a substantial degree of autonomy in the 19th century (Canada, Australia, New Zealand, South Africa).

I assume that governments set trade policy simultaneously, and look for the Nash equilibrium of this non-cooperative game. Because of the stark trade policy, the game admits a number of prisoner dilemma type of equilibria. Since it is imaginable that governments could easily coordinate out of these,\(^{14}\) I restrict my attention to the subset of coalition-proof Nash equilibria of

\(^{14}\)For example, it is a Nash equilibrium of the game that all countries refuse to trade with all other other countries, and no trade occurs. In such a situation, it would be strictly Pareto improving for any two governments to admit trade with each other’s country (except for two identical countries, such as \( B \) and \( F \) are if \( b = f \); in that case gains from trade would be zero for both countries).
the game. My first result is contained in Proposition 1:¹⁵

**Proposition 1.** If all countries set trade policy independently, there is trade between C and both B and F (and world-wide free trade realizes) at any coalition-proof Nash equilibrium.

An intuitive explanation of Proposition 1 can be readily found by investigating the preferences of government over alternative trade outcomes. I begin by deriving the indirect utility (welfare) of country J. In an endowment economy, this can be expressed as a function of prices only:

\[ v^J(p) = \frac{1 + py^J}{2\sqrt{p}}. \tag{5} \]

I show in Appendix A that the above function is strictly increasing in \( p \) for \( p > \frac{1}{y^J} \), strictly decreasing in \( p \) if \( p < \frac{1}{y^J} \). Quite intuitively, a country that exports manufactures would like to face the highest possible price of manufactures, while a country that imports them would like just the opposite. From our earlier discussion, we know that B and C are always exporters, while C is always an importer. It follows that the former two countries’ first best is achieved for trade conditions that give the highest possible \( p \), while C’s first best is achieved for conditions that give the lowest possible \( p \).

Now consider how a coalition proof Nash equilibrium can be formed in the model’s specific trading world. Because I have assumed that the key difference in comparative advantage lies between B and F, on one side, and C on the other, there is a natural tendency for C to be part of any trading block that emerges in equilibrium. After all, the really valuable trade is that between Europe and the New World, so it cannot be optimal for B and F to only trade between themselves - they would both benefit from opening up to C as well! But if C must be trading in equilibrium, then we know that both B and F must be trading with it. This is because the price of imported manufactures is lowest for C when competition between B and F unfolds, and this country would always stimulate this by opening to direct trade with both exporters.

Thus, Proposition 1 suggests that, in the counterfactual scenario to the centrally-administered European empires of the 16th-18th century, the New World would have thrown its market open.

¹⁵All proofs to the propositions are provided in Appendix B.
to competition between European powers. In other words, had Asian spice producers remained independent, they would have sold their products to the highest offerer between the Dutch and English East India Company; and the hypothetical independent settlers’ communities of the Americas would have placed their silver, sugar and tobacco directly onto the most interested European consumers, while sourcing their imports of manufactures from the most efficient European producers.

Having considered how the trade policy equilibrium would look like in the absence of imperialism, I next consider how such an equilibrium would change if $C$ falls under the imperial control of either $B$ or $F$.

### 2.3 Trade policy under imperialism

In this section, I simply assume that $C$ has fallen under the control of an imperial power, say $B$,$^{16}$ and study how this affects the equilibrium trade policy. In the next section, I will turn to the process through which $C$ may fall under imperial domination.

I model the establishment of imperial power as a transfer of policy-setting power from the government of $C$ to the government of $B$. However given the restrictions that I have imposed on the policy space, this amounts to assuming that all that $B$ gains from acquiring imperial control over is $C$ is control of its trade policy. Implicitly, I am assuming that other sources of colonial extraction - such as the appropriation of colonial land by European residents or the wages that European administrators earned while in the service in the colonies - were small relative to the gains from controlling colonial trade policy. I return to the plausibility of this assumption in section 4.

Suppose then that the “British Empire” has been created (call this $BC$), and the government of $B$ controls trade policy in both $B$ and $C$. Taking for granted that it will always allow trade between the two member component of the empire - that this is optimal is, again, a consequence of the assumed high relative importance of colonial trade - there are really two options that the government is left with: to keep the colonial market open to $F$’s trade, or rather to close it down.

$^{16}$Given the symmetry built in in the model, the choice of a specific imperial power does not imply a loss of generality.
Notice that, because there are no transport costs and $F$’s exports are perfect substitutes of $B$’s exports, this choice is substantially identical to that between keeping the whole empire open to $F$’s trade and closing it down. This is because $B$ could not possibly prohibit $F$’s manufactures on $C$’s market, while admitting them on $B$’s market: even if a prohibition to transship these manufactures through $B$ could be enforced, perfect substitutability would still depress the price to its level in the world-wide free trade equilibrium. In other words, such an arrangement would provide domestic manufacturers with a larger share of the colonial market, but it would cost them an equal and offsetting loss on the domestic market. This equivalence will have meaningful consequences for the economic impact of imperialism, as we shall see in more details below. The next proposition looks at the optimal choice of trade policy for an imperial power:

**Proposition 2.** An imperial power $J$ always prohibits trade between $C$ and $-J$, while allowing trade between $C$ and itself.

**Corollary.** In order to prohibit trade between $C$ and $-J$, the imperial power must also prohibit trade between itself and $-J$.

A good intuition for why an imperial power always restricts access to its colony can again be found by using the indirect utility function derived in 5. Because it is an exporter of manufactures (or services), $B$’s utility is highest when the price of $y$ is as high as possible, or, equivalently, when the price of raw materials is as low as possible. Because the main market for $y$ is in $C$ in this model, and $F$ compete with $B$ for this market, we may expect that excluding $F$ from the colonial market must yield $B$ a higher price. At the same time, we may expect that such a move must give $F$ a lower price. This if confirmed by the following picture, that compares the price of manufactures obtained by $B$ and $F$ when trade only takes place between the British Empire to the price in the world-wide free trade equilibrium.

Putting things together, it is clear that the reason why an imperial power always prohibits the external trade of her colonies is to improve her terms of trade vis-a-vis the colony. In other

---

17 A departure from the world-wide free trade equilibrium would only occur if $B$ was not able to fully supply the colonial market at world-wide free trade prices. This scenario can be ruled out in our parameter space: even if $b = 1$ and $f = 2$ ($B$ is little industrialized and the world price of manufactures is relatively low, leading to a relatively high colonial demand), $B$ would be able to supply $C$ with its required 0.5 of imports.
words, the imperial power is able to import a larger amount of raw materials for any amount of manufactures or services that she exports, thus capturing a larger share of the value of colonial trade. Of course, such an arrangement damages not only the other imperial power, who is excluded from colonial trade, but also the colony, who would have obtained a better value for its products in the world-wide free trade.\footnote{Notice that the colony pays a price $p^{BC} > p^W$ for her imports.}

Thus, Proposition 2 suggests that one rationale behind the setting up of centralized empires in the 16th-18th century was the desire to exclude European competitors from the lucrative profits that could be made on colonial trade. The large impact that such profits had on some of the countries of Atlantic Europe has been recently emphasized by Acemoglu, Johnson and Robinson (2005), who have shown that a key determinant of the English and Dutch merchants’ ascent to political power was the increased economic power stemming from high profits on colonial trade. It has also been suggested that high profits from colonial trade were particularly cherished by European rulers, at a time in which competition for power in Europe was at its highest and state revenues hard to raise (Ferguson, 2011). My model suggests that the establishments of centrally controlled empires was the institutional arrangement that countries of Atlantic Europe devised to maximise the inflow of such profit, to the detriment of both European competitors and of colonies.

The corollary to Proposition 2 has also potentially far-reaching implications. I have already discussed the intuition for this result, but this could be rephrased as follows. We know from Proposition 2 that $B$’s trade policy is geared towards increasing the price that $B$’s exporters
receive on C’s market, thus redistributing a share of colonial wealth from the colony to the imperial power. However higher export price must result in a higher price of y in the imperial power, which will create a pressure for domestic consumers to substitute goods produced domestically for ones imported from F. Unless these imports are restricted, this would result in a collapse of prices in B, ultimately transferring a fair share of the gains from colonial trade to F. These results imply that the conquest and ring-fencing of the New World markets in the 16th-18th century may have contributed to making European imperial powers anxious to protect themselves from one another’s exports, and thus eventually to the erection of strong trade barriers in Europe during the age of mercantilism (1650-1770).

This section has suggested one potential driver for the creation of centralized empires in the 16th-18th century, or in what I have labelled the phase I of European imperialism. But how was the actual creation of these empires determined in the strategic interaction of the countries of Atlantic Europe? And what determined the fluctuations in European imperialism that took place in phases II and III? To these questions I turn in the next two sections.

2.4 Gains from imperialism

The results of the previous two sections suggest what are the gains from empire in this model. If C remains independent, there is always competition between B and F, and both receive the world-wide free trade equilibrium price on their exports. If C is subjected to imperial control, on the contrary, the imperial power closes the colony down to trade with her competitor, thus obtaining a higher price for herself ($p^{BC} > p^W$), and a lower price for her competitor ($p^F < p^W$).

It follows that there are two distinct sources of benefit from empire. On one hand, to set up an empire gives the imperial power a larger share of the gains from colonial trade, relative to what she would have obtained in the world-wide free trade equilibrium. We call this the “offensive” component of the gains from empire, $\pi^J_\alpha$. For imperial power $J \in \{B, F\}$, this is equal to:

$$\pi^J_\alpha(b, f) = v^J[p^{JC}(b, f)] - v^J[p^W(b, f)],$$

(6)
where I have recognized that prices (and thus the gains from imperialism) are a function of \( b \) and \( f \) only.

On the other hand, to set up an empire has also the benefit of preventing the imperial power’s competitor from setting an empire of his own. We call this the “defensive” component of the gains from empire, \( \pi^d_i \):

\[
\pi^d_i(b, f) = v^f[p^W(b, f)] - v^f[p^I(b, f)].
\] (7)

We also define the total benefit from empire as the same of its offensive and defensive components:

\[
\Pi^I(b, f) = \pi^o_i(b, f) + \pi^d_i(b, f).
\] (8)

We now want to find out how \( \pi^o_i(b, f) \), \( \pi^d_i(b, f) \) and \( \Pi^I(b, f) \) depend on the parameters \( b \) and \( f \). In what follows, I continue to refer to the three scenarios discussed in figures 2-4. To recap, these are, \( b = f = k \), with \( k \) ranging between 1 and 2; \( b \) ranging between 1 and 2 with \( f \) fixed at 1; and \( f \) ranging between 1 and 2 with \( B \) fixed at 2. To acknowledge that \( B \) and \( F \) are identical in the first scenario, I drop the superscript \( J \) in this case. Moreover, I express the gains as function of the common \( k \). The following Lemma contains my results:

**Lemma 1.** The gains from imperialism depend on the endowment parameters as follows:

- **In the first scenario**, both \( \pi^o_i(k) \) and \( \pi^d_i(k) \) are strictly increasing in \( k \).

- **In the second scenario**, \( \pi^o_i(b, 1) \) is strictly decreasing in \( b \), while \( \pi^F_i(b, 1) \) is strictly increasing. On the contrary, \( \pi^d_i(b, 1) \) is strictly increasing in \( b \), while \( \pi^F_i(b, 1) \) is strictly decreasing. This results in \( \Pi^F(b, 1) \) being strictly increasing in \( b \), and in \( \Pi^F(b, 1) \) being constant.

- **In the third scenario**, \( \pi^o_i(2, f) \) is strictly increasing in \( f \), while \( \pi^F_i(2, f) \) is first increasing then decreasing. On the contrary, \( \pi^d_i(2, f) \) is strictly decreasing in \( f \), while \( \pi^F_i(2, f) \) is strictly increasing. This results in \( \Pi^F(2, f) \) being constant, and in \( \Pi^F(2, f) \) being strictly increasing in \( f \).
The pattern described in Lemma 1 may seem chaotic at first, but it is actually easy to rationalize intuitively. In the first scenario (depicted in the first column of figure 6), I consider the possibility that $B$ and $F$ “industrialize” (an increase in $k$) at the same time. As this happens, the two countries’ demand for raw materials rise, and so does competition for colonial trade. Not surprisingly, this is associated with both a higher gain from excluding one’s competitor from the colonial market - a higher $\pi_o(k)$ - and a higher cost from being excluded - a higher $\pi_d(k)$.

In the second scenario (second column of figure 6), I am looking at the case in which $B$ industrializes, while $F$ remains at her starting point. As this happens, total demand for colonial materials increases, but this is now entirely driven by $B$ (demand in $F$ actually falls). This implies that, in the world-wide free trade, $B$’s share of colonial trade increases. Figure 2 shows that it actually converges to one as $b$ approaches 2. In this scenario, $B$’s offensive gains from imperialism decline, since the increase in market share that the closure of the colonial market is associated with becomes more and more modest. On the contrary, the cost to $B$ from being excluded from the colonial market is higher the higher is her stake in it. Not surprisingly, then, $B$’s defensive gain is increasing in $b$. The opposite pattern holds for $F$: this country has a decreasing share of $C$’s market as $b$ increases, and her gains from closing the colonial market (avoiding being excluded) is then naturally increasing (decreasing). Notice that the total gain from imperialism is increasing for $F$, constant for $B$: the latter being due to the fact that neither $p^{FC}$ nor $p^F$ depend on $b$.

Finally, in the third scenario (third column of figure 6), $F$ is “catching up” to an industrialized $B$. Again, total demand for raw materials increases, but this is now driven by higher demand in $F$ (which offsets a lower demand in $B$). As $F$’s market share in $C$ increases, her benefit from excluding her competitor (eventually) declines, while her interest in preventing to be excluded rises. On the contrary, $B$’s offensive interests increases as she loses market share, while her defensive interest decreases. On the the whole, $F$’s gain from imperialism increases, while it is now $B$’s gain to be constant (since $p^{BC}$ and $p^B$ do not depend on $f$).

On the whole, these results suggest that a balanced pattern of industrial development makes all potential imperial powers more interested in their own, and more concerned about the others’, offensive imperialism. On the contrary, unbalanced industrial development - be it in the form
of industrial taking off, or of industrial catching up - shifts the focus of the fastest growing imperial power from the offensive to the defensive, while doing exactly the opposite to her rival. In the next section, I look at how these fluctuations in countries’ interests affect their strategic interaction, ultimately determining whether C is annexed or not.

2.5 Costs of imperialism and equilibrium

Having defined the benefit from empire, I next consider its costs. I assume that these stem from two separate sources. On one hand, there is a cost $k_A$ from establishing and maintaining administrative control over the colony. Because the colony incurs a positive cost from being part of an empire, a natural interpretation of this cost is that it originates from the need to avoid or suppress colonial rebellions. Alternatively, it could originate from the inefficiency generated by centralized control, particularly in a world in which long-distance communications are very slow.\textsuperscript{19} For simplicity, I take the cost of administrative control as exogenously given.

\textsuperscript{19}One would imagine that the colony has its own public revenues that the imperial power can use to pay its administration costs; however such revenues may not be enough to pay for the inefficiencies of centralized control, thus demanding a net transfer of resources from the imperial power to the colony.
On the other hand, \( B \) and \( C \) may be unable to annex \( C \) before fighting a costly war with each other. I assume that each country \( J \in \{B, F\} \) must decide whether she is “internationally active” \( (a^J = 1) \) or not. By internationally active, I mean that they deploy the sort of military (naval) strength that is needed to establish an area of influence around \( C \). In order for a country to have the option to annex \( C \) (at a cost \( k^A \)), she must first become internationally active. However if both countries are internationally active, a war between them takes place, following to which both countries report an exogenous cost \( k^M \). There is always a winner to the war, and this country receives the option to annex \( C \). I assume that the probability that each country wins the war is increasing in her relative level of industrial development. This may captures the fact that highly industrialized countries are better positioned to produce sophisticated military equipment. Alternatively - referring in particular to the early centuries of European imperialism - it may captures the fact that largest exporters are likely to have a larger merchant fleet, and the latter may be instrumental to winning a naval war.\(^{20}\) The probability that \( B \) wins the war, \( q \), is then captured by the following contest function:

\[
q = \frac{b}{b+f}.
\]  

(9)

Summing up, the timing of the game is as follows:

1. \( C \) is discovered, and the possibility to trade with it emerges;

2. \( B \) and \( F \) simultaneously and independently decide whether to be internationally active or not. If they both are, a war ensues which entails a cost \( k^M \) to each country, and that is won by \( B \) with probability \( q \).

3. A country that is alone in being internationally active has the option of annexing \( C \) at a cost \( k^A \). Alternatively, the winner of the war has that option.

4. All countries set trade policy simultaneously, with trade policy in \( C \) being determined by

-----

\(^{20}\)Howard (1976, p. 41) argues that merchants ships were not much less effective than warships in the 17th and early 18th century, since they normally carried guns and had a comparable gun power. In fact, the merchants played a key role in the colonial appendix to many European wars in this period (Ibid, p. 42 and 48; Ferguson, 2011, p. 38).
her imperial power if she has been annexed.

5. Trade take place, payoffs realize.

I solve the game using backward induction. From Propositions 1 and 2, we know that, from the standpoint of a country who has the option of annexing $C$ in period 3, the gains from empire consist in being able to exclude her rival from the colonial market. That is, the gains from empire are purely offensive in this period, since the option to annex $C$ cannot belong to more than one country. Such offensive gains from empire are then compared with the exogenous cost of administering the colony, $k^A$. To focus on the case of interest - that is, one in which offensive imperialism is strictly profitable for two countries at the same level of industrial development - I assume the following:

**Assumption 1.** $\pi_o(1) > k^A$.

In period 2, countries must make the more complex decision on whether to be internationally active or not. At this early stage in which the imperial race is still open, there are two potential gains from being internationally active. On one hand, it may give a country the option to annex the colony, and exclude competitors from her trade. This option is not necessarily a profitable one, since the cost of administering the colony may well outweighs the value of monopolizing her trade. On the other hand, being internationally active may prevent colonial trade from being monopolized by the other country, either by successfully deterring the other country from being internationally active, or by engaging it in a war. Both of these gains, as well as the probability that the country wins the war, depend on the structure of endowments as described in Lemma 1 and equation 9. There is then an obvious cost of being internationally active, which results from the cost of war. Again, I introduce a parametric assumption that focuses attention on the case of interest - one in which offensive imperialism is strictly profitable for two countries at the same level of industrial development, even if this implies going to war against each other:

**Assumption 2.** $\frac{1}{2} [\Pi(1) - k^A] > k^M$.

The equilibrium actions of the two countries in periods 2 and 3 is described in the following proposition. I continue to refer to the three scenarios for the value of parameters that I described at length in the previous section:
Proposition 3. In equilibrium:

- If parameters are as in the first scenario, both $B$ and $F$ are internationally active, a war is fought, and the winner annexes $C$;

- If parameters are as in the second scenario, there exists $b' \in (1, 2]$ and $b'' \in (1, 2)$ such that:
  - If $b < b', b''$ both $B$ and $F$ are internationally active, a war is fought, and the winner annexes $C$;
  - If $b' \leq b < b''$, only $B$ is internationally active, and she annexes $C$.
  - If $b'' \leq b < b'$, $B$ is always internationally active, while $F$ may or may not be; however $C$ is only annexed if $F$ is active, and wins the war;
  - If $b', b'' \leq b$ only $B$ is internationally active, and she does not annex $C$;

- If parameters are as in the third scenario, there exist $f' \in (1, 2)$ and $f'' \in [1, 2)$ such that:
  - If $f < f', f''$, only $B$ is internationally active, and she does not annex $C$;
  - If $f'' \leq f < f'$, only $B$ is internationally active, and she annexes $C$.
  - If $f' \leq f < f''$, $B$ is always internationally active, while $F$ may or may not be; however $C$ is only annexed if $F$ is active, and wins the war;
  - If $f', f'' \leq f$ both $B$ and $F$ are internationally active, a war is fought, and the winner annexes $C$;

Proposition 3 outlines a very diverse pattern of international activism, war, and imperial expansion. For given costs of colonial administration and war (which determine the thresholds $b'$ and $b''$), such pattern is determined uniquely by the structure of endowments, parameters $b$ and $f$. The first scenario of interest is one in which $b$ and $f$ are identical, and take value in $[1, 2]$. This considers - if in a very stylized way - how the international conduct of $B$ and $F$ is affected by their simultaneous, and symmetric, industrial development. That at $b = f = 1$ the equilibrium is as described, it follows directly from Assumptions 1 and 2. There, I assumed that monopolizing colonial trade is valuable enough relative to the cost of colonial administration;
and that doing so by, at the same time, avoiding that colonial trade is monopolized by the other country, is valuable enough relative to the cost of colonial administration and of war. It is then obvious that, in equilibrium, both countries are active, there is a war, and the winner annexes $C$. The contribution of Proposition 3 is to say that the equilibrium remains the same as the two countries become more industrialized. The reason for this result is simple: simultaneous industrialization makes colonial trade even more valuable, without affecting the two countries’ market share in it. It follows that both market shares become more valuable, which in turn implies that both the offensive and defensive gains from imperialism becomes larger (as shown in the first column of figure 6). Since the balance of military power is not affected by symmetric industrialization, the same identical factors that support the imperial equilibrium for $b = f = 1$ are there - in fact, even stronger - at higher levels of industrialization.

The second scenario also includes $b = f = 1$ as a starting point, but considers the case in which only $B$ industrializes ($b$ increasing from 1 to 2), while $F$ lags behind ($f$ fixed at 1). Here, there are two distinct effects of this industrial taking-off on equilibrium imperialism. On one hand, industrial taking off gives $B$ an advantage in the international arena, by making it more military powerful. In fact, there may be a level of industrial development in our parameter space - the threshold $b'$ - above which $F$ surrenders any international activity to $B$. On the other hand, industrial taking off, despite making colonial trade more valuable, increases $B$’s market share in it. This decreases the offensive benefit from imperialism all the way down to zero (second column of figure 6), implying that there must be a second threshold in our parameter space - $b''$ - above which $B$, if presented with the possibility to annex $C$ at a cost $k^A$, declines. At the same time, $B$’s defensive gains from imperialism increases with $b$, since a more valuable market share is more costly to loose. Also, $F$’s imperial incentives move in the opposite directions as $B$’s.

It follows that equilibrium imperialism depends on the stage of $B$’s industrial take off. If this has just started ($b < b', b''$), the imbalance of military power and $B$’s market share are still not too large. In this case, $F$ remains internationally active, in part to try to prevent $B$ from annexing $C$. With both countries active internationally and intentioned to annex $C$, there is always war and empire building; the only difference relative to the first scenario is that the stronger country, $B$, has now a higher (and increasing) probability to succeed (see figure 24).
7). On the opposite side, if industrial take-off is largely achieved \((b', b'' \leq b)\), the imbalance of military power and \(B\)'s market share are very large. In this case, \(F\) withdraws from international activism, partly in recognition of her military inferiority, but partly also in expectation that \(B\)'s international activism will not result in empire building. This expectation is, of course, correct: being left as the only internationally active country, \(B\) opts for not annexing \(C\). Notice that \(B\) must be internationally active in equilibrium: if she didn’t, nothing would be deterring \(F\) from becoming active and annexing \(C\) - something that, given \(F\)'s relative economic weakness, would always be optimal. For intermediate stages of industrial take-off, there are two cases. If \(b' \leq b < b''\) (the case represented in figure 7) \(B\) is both military dominant and intentioned to annex \(C\). This guarantees that this country is uncontested in equilibrium, and always annexes \(C\). If \(b'' < b < b'\), the imbalance of military power is not strong enough to guarantee that \(F\) withdraws from international activism.\(^{21}\) However independently on whether \(F\) is active, the probability that \(C\) is annexed is lower than for \(b' < b', b''\), since \(B\) is not intentioned to annex it.

![Figure 7: Probability of empire as a function of endowments, case \(b' < b''\) and \(f < f''\).](image)

The third scenario represents a largely symmetric situation, in which \(F\) catches up \((f\) increasing from 1 to 2\), while \(B\) has exhausted her growth period \((b\) fixed at 2\). Again, there are two key thresholds. The first, \(f'\), determines whether industrial catching up has gone far enough to reduce the imbalance of military power between \(B\) and \(F\), that the latter country is willing to return to international activism \((f \geq f')\). The second, \(f''\), determines whether industrial

\(^{21}\)The condition \(b \geq b'\) is sufficient but not necessary for \(F\) to withdraw from international activism. I am exploiting the fact that \(\Pi^F(b, 1)\) is constant to define the stage of industrial take-off above which \(F\) withdraws from international activism, even if \(B\) is intentioned to annex \(C\). If \(b'' < b'\), \(F\) has even stronger reasons to withdraw when \(b' < b\), since \(F\) is not intentioned to annex \(C\); when \(b'' < b < b'\), however, it may still be optimal for \(F\) to withdraw, despite a smaller imbalance of military power, in recognition that \(C\) does not intend to annex \(C\).
catching up has gone far enough that $B$, having lost a big chunk of her colonial market share, is willing to return to offensive imperialism ($f \geq f''$) to promote her economic interests abroad.

When neither conditions has yet realized, ($f < f', f''$), the equilibrium remains the same as for $b', b'' \geq b$ in the previous scenario: although $F$ is on the road to industrial recovery, $B$’s military and economic power is still overwhelming enough for this country to act as a stabilizer of international relations. On the opposite side of the spectrum (when $f', f'' \leq f$) the equilibrium falls back to the that prevailing at lower levels of industrial development, since the two countries are now very similar and industrial development has made colonial trade even more valuable.

In the middle, there are two cases. If $f'' \leq f < f'$, the imbalance of power is still large, but $B$’s economic supremacy has been largely lost. We then return to an equilibrium in which $B$ is uncontested in the international arena, and uses this position to secure markets abroad. If $f' \leq f < f''$, it is the other way round. Now $F$ may want to stand up in the international arena, but his motivation is tempered by the fact that $B$ still does not have an offensive purpose at this stage. Thus, the probability of $C$ being annexed can still be zero, but it is actually positive whenever $F$ is internationally active (figure 7 represents this case, with $F$ not standing up in the international arena).

In this section, I have considered a world in which trade between $B$ and $F$ plays essentially no role, since the goods and services that these two industrial countries export are perfect substitutes, and their comparative advantage is in exchanging them with the raw materials-abundant part of the world. While this has allowed me to focus very sharply on the role of colonial trade in the early phases of European imperialism, trade between the industrial nations has always played an important role in world trade. Furthermore, its relative importance has risen sharply in the 20th century, the beginning of this rise roughly coinciding with the coming of decolonization. It thus seems to make sense to consider a more general model of trade, both to gauge how robust the results of this section are and to see if I can say anything about the economic forces that led to decolonization. To this task I now turn.
3 Generalizations

Suppose now that, rather than being homogeneous, good \( y \) comes in two differentiated varieties, \( y_1 \) and \( y_2 \). Preferences are still described by a Cobb-Douglas function, but \( y \)'s input is now a CES aggregator:

\[
U = x^{\frac{1}{\sigma}} \left( \frac{x^{\frac{1}{\sigma}}}{y_1^{\frac{1}{\sigma}} + y_2^{\frac{1}{\sigma}}} \right)^{\frac{\sigma}{\sigma-1}}
\]

(10)

where \( \sigma \) is the elasticity of substitutions between varieties of \( y \) and is assumed to be defined in \((1, \infty)\).

The endowments of \( x \) have not changed, but endowments of \( y \) reflect the fact that each country produces different varieties of this good:

\[
\begin{align*}
    y_1^B &= b & y_1^F &= 0 & y_1^C &= 0 \\
    y_2^B &= 0 & y_2^F &= f & y_2^C &= 0
\end{align*}
\]

where \( b \in [1, 2] \) and \( f \in [1, 2] \) as before. It is important to notice two things about this modified model. First, it contains the previous model, in that it converges to it as \( \sigma \to \infty \). Second, the model can be seen as a three-country version - if a stylized one - of the nested Heckscher-Ohlin and “New Trade” model, first introduced by Krugman (1981). On one hand, the assumption of different endowments of \( x \) and \( y \) across countries captures the different types of factors needed to produce these two goods, and their distribution across countries. This is line with standard Heckscher-Ohlin theory. On the other hand, the assumption that alternative varieties of \( y \) are produced only in different countries is harder to reconcile with a constant-returns to scale HO world, because it is imaginable that similar factors of production will be used in the production of the varieties. Instead, this assumption makes the model’s predictions akin to those of a model with increasing returns to scale, in which each variety can only be produced in one country in the presence of trade. In fact, to change \( b \) or \( f \) in this model has essentially the same impact on welfare as to change the endowments of the factors of production that determine the number of varieties produced in each country in a fully specified New Trade model.
After normalizing the price of $x$ to 1, there are now two prices that we need to solve for - one for each variety of $y$. The equilibrium price of each variety when free trade realizes within a subset $M$ of countries can be found to equal:

$$p_i^M = \frac{x^M}{(y_i^M)^{\frac{1}{\sigma}} \left[ (y_1^M)^{\frac{\sigma-1}{\sigma}} + (y_2^M)^{\frac{\sigma-1}{\sigma}} \right]}$$

(11)

where $y_i^M \equiv \sum_{j \in M} y_{ij}$. The indirect utility of a country facing prices $p_1$ and $p_2$ is now:

$$v^J(p_1, p_2) = \frac{1 + p_1 y_{1J} + p_2 y_{2J}}{2\sqrt{P}}$$

(12)

where $P \equiv (p_1^{1-\sigma} + p_2^{1-\sigma})^{\frac{1}{1-\sigma}}$ is the CES price index.$^{22}$

Just as before, in the world-wide free trade equilibrium, $B$ and $F$ compete to exchange manufactures for raw materials with $C$. There are, however, two important differences. First, $B$ and $F$ compete to sell varieties of manufactures that are only imperfect substitutes of each other. Second, trade in manufactures occurs between the two countries. This pattern is illustrated in figure 8. For brevity, I from now onwards only deal with the previous sections’ “first scenario” ($k \equiv b = f$, with $k$ taking value in $[1, 2]$). The three panels of figure 8 plot trade flows in the $x$, $y_1$ and $y_2$ industries respectively.

Figure 8: Trade at the world-wide free trade equilibrium.

Again, industrialization in $B$ and $F$ maps into a constant export flows of raw materials from

\footnote{This can be seen as the price of the CES aggregator, $y \equiv \left( \frac{y_1}{y_1^{\frac{\sigma-1}{\sigma}} + y_2^{\frac{\sigma-1}{\sigma}}} \right)^{\frac{\sigma}{\sigma-1}}$.}
C, but in a declining relative price of manufactures (first panel of figure 9). These two facts put together imply that the volume of manufactures exported to C is increasing, and so is the volume of manufactures that B and F exchange with each others. The second panel of figure 9 shows that both the share of total trade to GDP, and the share of imports of manufactures to GDP, are constant as GDP grows in the two industrial countries.

![Figure 9: Relative price of manufactures and trade to GDP ratios.](image)

Prices are evaluated at $\sigma = 5$. $M^J$ is country J’s imports plus exports, and both $M^J$ and $GDP^J$ are expressed in terms of the numeraire good.

As in the previous section, I want to compare the trade policy equilibrium that realizes when each country sets trade policy independently, to the one that realizes when B (say), being the imperial master of C, sets trade policy for himself and for C as well. It turns out that Proposition 1 still holds in the modified model. As for Proposition 2, this becomes:

**Proposition 2-bis.** There exists $\overline{\sigma} \in (1, \infty)$ such that, if $\sigma \geq \overline{\sigma}$, an imperial power J always prohibits trade between C and $-J$, while allowing trade between C and herself. If $\sigma < \overline{\sigma}$, on the contrary, the imperial power always allows trade between C and $-J$.

Proposition 2-bis qualifies the results of Proposition 2 (for the first parameter scenario) by revealing that the results there stated only holds for the case in which $\sigma$ is high enough (recall that we have implicitly assumed $\sigma = \infty$ in section 2). The intuition for this result is as follows. On one hand, restricting extra-imperial trade gives the imperial power the same benefit as before: it increases the price that the imperial power obtains on her exports of manufactures (or services). Since neither $p^W$ nor $p^{BC}$ depend on $\sigma$ in this scenario, this benefit is exactly the

---

23 Given symmetry, the price of each variety is the same as that of the homogenous good in the simpler model, while the lower price index reflects the gains from differentiation.

24 This is evident by plugging the relevant endowments in equation 11.
same as in the simple model. On the other hand, restricting extra-imperial trade increases the price that imperial consumers have to pay for manufactures. This increase is higher when $\sigma$ is high, reflecting the fact that it is more welfare decreasing to lose access to a foreign supply of manufactures when these are highly differentiated, than when they are homogeneous. Because of this higher cost, restricting extra-imperial trade is less beneficial the more differentiated are manufactures, and becomes a net loss for $\sigma$ high enough. When this is the case, an imperial power always finds it optimal to open extra-imperial trade, thus obtaining the same payoff she would have obtained had C remained independent.

This result is very similar to the one found, in a different context, by Krugman (1981). That paper investigates the conditions under which it is welfare-improving for owners of a scarce factor of production (used to produce differentiated manufactures) to open up to international trade, in the presence of both inter and intra-industry trade. A key result of the paper is that, for any given relative importance of intra-industry trade, there exists a threshold for the degree of differentiation of manufactures above which owners of the scarce factor gain from opening up to trade. Thus, as we move away from a traditional Heckscher-Ohlin world (minimum degree of differentiation) to a world with highly differentiated manufactures, the owners of the scarce factor switch from being against trade to being in favor. With the notable difference that our single country is now an empire, this is precisely the result contained in Proposition 2-bis. As $\sigma$ decreases, owners of the scarce factor within the empire - that is, owners of manufactures - switch from being against opening up the empire to external trade, to being in favor.

Since the two imperial powers can be expected to report a net loss from empire if $\sigma > \sigma$ (after taking the cost of administration, $k^A$, into account), one would expect the equilibrium of their strategic interaction to be characterized by a lack of empire building. This is confirmed by the next proposition:

**Proposition 4.** With parameters as in the first scenario and $\sigma < \sigma$, neither or only one of B and F is internationally active, and no one annexes C.

Proposition 4 confirms that if manufactures are sufficiently heterogeneous, so that they have an important market in other manufacture exporters, two potential imperial powers reject empire building as a way to promote international trade. Notice that, while it is possible that either of
them is still active internationally, this is in the context of an equilibrium in which the inactive
country trusts the active country not to take advantage of such position. This is important
because it suggests that, were there any additional reasons for countries to become internationally
active, these could be pursued without the potential complication of triggering an empire-building
race.

4 The rise and evolution of European imperialism, 1492-1938

The central questions addressed in this paper are: Why did Europeans chose to build centrally
administered empires in the 16th-18th centuries? What explains their evolution and final col-
lapse, as described in Figure 1? And what was the impact of imperialism on contemporary trade
relations between the European powers? In the previous two sections, I have sketched the main
lines of an argument providing a possible answer to these questions. I now turn to describing
the argument in greater detail. I begin by considering the history of imperialism from its rise in
1492, to the early decades of the 20th century (phases I-III in figure 1). In the next section, I
will then consider its final collapse in the 20th century (phase IV).

The main lines of the argument for the period 1492-1938 can be summarized as follows:

1. The great discoveries of the late 15th century opened up a “new” intercontinental trade in
foodstuff and raw materials for manufactures and services, that would remain important
for the European economy for more than four centuries.

2. The institutional set-up that this led to was one in which Europe established and main-
tained direct imperial control over the new trade partners - as opposed to accepting them
as independent counterparts.

3. One key role of such institutions was that it allowed imperial powers to secure a dispropor-
tionate share of the value of colonial trade, relative to what they would have achieved had
the foreign territories remained independent, or fallen under the control of other European
countries (Propositions 1 and 2).
4. A high value of colonial trade, combined with close economic competition between the economies of Atlantic Europe, led to an equilibrium in the 16th-18th century in which all countries of Atlantic Europe were military active on the high seas, frequent wars ensued, and winners annexed as much overseas territory as possible (first point of Proposition 3).

5. In 1830-1880, the emergence of Britain as a clear industrial leader created a new equilibrium in which Britain was disproportionally active on the high seas, but didn’t annex new overseas territories for the sake of monopolizing their trade (second point of Proposition 3). British imperialism of this period was rather informal, or, when it was formal, it pursued goals other than direct trade benefits.

6. The economic and military decline of Britain relative to continental Europe and the US may have implied the collapse of such an equilibrium after 1880, explaining the late-19th century wave of imperialism (third point of Proposition 3).

7. The creation of colonial empires created large “domestic” markets for imperial manufactures, and ones in which consumers, being partly colonial, had a lesser weight in government objectives. This created a pressure for higher protection of domestic markets, and may contribute to explaining the high tariffs existing in Europe in periods of rapid imperial expansion (e.g. the 18th century; 1880-1938).

In what follows, I describe the argument in more detail by expanding on each of the above-listed points.

1 - The rise of colonial trade

With the discovery of the Americas by Christopher Columbus and the sea way to the Indies by Vasco de Gama in the late 15th century, Europeans gained access to a completely new trade in the West, and obtained a drastic abatement in transport costs in the old and valuable trade with the East. The result was the rise of a “new” intercontinental trade between Europe and the rest of the world. With some major exceptions, this was mostly an exchange of foodstuff and raw

\textsuperscript{25}A notable exception was the import of Indian textiles, which was prominent from the late 17th to the early 19th century.
materials produced in the rest of the world for manufactures and services produced in Europe. In particular, the import side was dominated by a few commodities, whose importance fluctuated over time. In the early centuries, these were mostly exotic foodstuff, while raw materials and temperate foodstuff became more important with the industrial revolution. These imports were paid for with a variety of manufactures, although silver was they key method of payment in the East in the 16th-18th century. However particularly in the early centuries, another key European export of Europe were the transportation and merchant services that were essential to connect overseas producers to European markets.

For their comparative advantage at oceanic navigations, the five countries of “Atlantic Europe” (Acemoglu et al., 2005) - Portugal, Spain, and later, England, France and the Netherlands - were the key early players in this trade, displacing older players that had dominated the over-land trade with Indies (such as the Ottoman empire and Venice). With the spreading of the industrial revolution from the second half of the 19th century, other European (Germany, Italy) and non European26 (the US, Japan) countries became important as well.

With some notable fluctuations, this intercontinental exchange remained a key component of the European economy for more than four centuries. Of course, before the 19th century high transportation costs made sure that trade was not a very high share of GDP for most European countries (with the exception of the Netherlands).27 However, intercontinental trade was still important for three reasons. First, it quickly became an important share of total trade, at least for the countries of Atlantic Europe. For example, figure 10 illustrates that British inter-continental exports rose to constitute around 50 % of British exports in the first three quarters of the 18th century. With some fluctuations and adjustments (the figure considers the US and Japan as “part of Europe” from 1870 onwards)28, it remained at this order of magnitude until

---

26Industrialization in the US changed the role that this country played in intercontinental trade, making it more similar to a European country than to a ROW exporter of foodstuff and raw materials.
27O’Rourke et al. (2008) report total trade (export + import) at the end of the 18th century to amount to 16% of GDP in Spain, 24% in Britain, and 20% in France (p. 30).
28The goal of the figure is to illustrate the relative importance of “potentially colonial” trade. By this, I mean trade, by and large in foodstuff/raw materials for manufactures/services, occurring with less developed nations. By the end of the 19th century, the US was an industrialized country more akin to an advanced European country than to a less developed nation. As for Japan, I add it in the same year as the US just for clarity. The basic message of the picture does not rely on when these two countries are added, provided that they are added by the end of the 1950s.
after World War II. O’Rourke et al. (2008) report figures (p. 28) illustrating that, at the end of the 18th century, inter-continental trade amounted to 38% of total trade in France, around 35% in Portugal and Spain, and 20% in the Netherlands, where however trade was a stunning 110% of GDP (p. 30).

Figure 10: Share of Europe in British exports, 1655-1985 (Europe + US + Japan from 1870 onwards).
Source: Davis (1954, 1962, 1979); Barbieri et al. (2008).

Second, inter-continental trade was of special importance for key groups in European societies. For example, Acemoglu et al. (2005) have argued that a key determinant of the English and Dutch merchants’ ascent to political power in the 17th and 18th century was the increased economic power stemming from high profits on inter-continental trade. It has also been argued that European rulers were particularly keen on securing the profits from the new trade, given the complications associated with imposing higher taxes on their domestic subjects (O’Rourke et al., 2008, p. 13-14), and the need to cover the military expenses associated with intense competition for political supremacy in Europe (Ferguson, 2011). Finally, it has been claimed that the profit from inter-continental trade were a key element that fostered economic growth in Europe, contributing to explaining its taking off in the 18th century (Allen, 2003, p. 432).

The most distinctive feature of the new trade was that it was, by and large colonial - that is, it occurred between two counterparts, one of which was institutionally subordinated to the other. To the description of these imperial institutions I now turn.
Most of the overseas populations that the Europeans got in touch with from the 16th century onwards were military weak relative to then, either already upon discovery, or at a later stage, following to sustained growth in Europe in the 18th and 19th century. This allowed for a progressive expansion of the Europeans, who eventually imposed political control over a large number of overseas territories. This expansion was often instrumental for intercontinental trade to take place, particularly in regions that were sparsely populated or little developed.

The bulk of European expansion took place in two distinct phases, 1492-1776 and 1880-1914 (phases I and III in Figure 1). Phase I began with the expansion of Spain and Portugal. In their “golden century” (the 16th century), these two countries annexed much of current-day Latin America and the Philippines, and greatly expanded Europe’s trade connections with the Indian Ocean. Expansion in the West continued in the 17th and 18th century, with Dutch, English and French annexing what was left of the Caribbean, and important parts of North America. In the East, territorial domination also became increasingly frequent with the arrival of the Dutch and the English. The landmark moment of this expansion was the establishment, after the battle of Plassey (1757), of de facto British control over Bengal. This laid the foundations for the expansion of the British Empire over the whole of India during the following 70 years. Phase III was much faster: between 1880 and 1914, a bunch of old and new imperial powers (mostly Britain, France, Germany and Italy, but also the US and Japan) expanded imperial domination over Africa, South-East Asia and parts of China.

The annexation of the new territories took very different forms across space and time. In the more sparsely populated territories of North America and Australia, annexation meant first of all colonization by large number of European settlers. In most other territories, administrative control was established through a military presence and a narrow cast of frequently turned-over bureaucrats and administrators (e.g. India, Ghana), possibly supported by small groups of settlers that remained minority in the local population (e.g. Spanish America, Algeria, Kenya). In many cases, this was done by allowing indigenous polities to continue to exist, under the agreement that they would become subordinated to the new rulers. In some other territories, a minority of Europeans was left to preside over largely imported populations of slaves (e.g. the
US South, the Caribbean islands).

One common characteristic of the imperial institutions set up by Europeans was that ultimate political power on a set of key policy issues - among which, major trade policy decisions - remained firmly in the hands of groups of people residing in the mother country. Normally, colonial government activities were carried out by governors coming from the mother country. These governors, which were frequently turned over, were appointed either by the government of the mother country or by the private companies that the government had entrusted with annexing and administering the overseas territory. While governors were responsible for the day-to-day administration of the colony (where an indigenous local government was not in place), decisions regarding on important matters were demanded to the mother country. More general, colonies were bound by the laws passed in the mother country.

To be sure - and as well known from the work of Acemoglu et Al. (2001) - the kind of local political institutions that Europeans implanted differed hugely from territory to territory. On one side, in colonies populated mostly by settlers, these brought with them some of the participatory political institutions that had been developed in Europe over the previous centuries. Thus, the British colonies of North America were normally endowed with local representative assemblies, with the power to advice governors on specific issues. On the opposite side, in colonies where a small minority of Europeans presided over a vast and subjugated labor force, more absolutists institutions were established. In both cases, however, ultimate political power rested with the mother country. Even in the above-mentioned case of British settlers colonies, the governor was in principle free to disregard colonial advice if instructed to do so.

That ultimate political power rested with the mother country was not, per se, bad for everyone residing in the colonies. For one thing, newly-established communities of settlers were often subsidized by the mother country, until they reached the capacity to pay for themselves. More in general, the mother country could provide, to at least some groups of colonial citizens, a variety of public goods, such as political stability, the rule of law, and a good trade and investment environment. For settlers, a key role played by the mother country was to guarantee colonial security, against both external and internal threats. External security involved protecting the

29This was, for example, the case of the British Australian colonies in the late 18th century.
colony from a potential take-over by another European country, which could imply expropriation of the existing settlers. Internal security was particularly relevant in plantation colonies, where settlers had only a limited capacity to keep vast populations of slaves under control.

Even for those colonial citizens that benefited from imperial rule, however, this had also important costs. This takes us to the discussion of the benefits from Empire, to which I now turn.

3 - Imperialism and the control of colonial trade

There was a number of ways through which a mother country could benefit from her empire. On one hand, she could herself benefit from some of the public goods that she provided for her colonial subjects. For example, this was true for the protection that she accorded against internal rebellions, since a low value of colonial wages was indirectly beneficial for Europeans as well. It was also true for the guarantee of domestic political stability, the promotion of a favorable environment for trade and investment, and the diffusion and consolidation of European cultural and religious values. On the other hand, there were a series of arrangements designed to transfer welfare from colonies to the mother country. Examples include the appropriation of colonial assets by absentee landlords, the high wages paid to colonial administrators, and the utilization of colonial public revenues for the achievement of broader imperial goals.

While these were all potentially important benefits, my focus in this paper is on one major benefit of empire, namely the capacity of mother country to regulate colonial trade. With a few major exceptions - among which the 19th century “imperialism of free trade” which I discuss at length below - imperial powers normally used this capacity to appropriate a disproportionate share of colonial trade for themselves.

In the early centuries, this was done by outrightly monopolizing colonial trade. There were two main ways in which this could happen. In some cases, the monopoly was awarded by the government of the mother country to private companies. The basic understanding underpinning these concessions was that the state would appropriate part of the companies’ profit, and that the companies would act as naval subcontractors against rival powers (Ferguson, 2011, p. 38).

I am not considering here the initial benefit accruing to settlers who grabbed colonial assets. This omission is appropriate in the current context, since my focus is on the benefit of empire for the mother country.
In facts, these companies were often the main engine behind the initial territorial expansion, and retained administrative rights over the colonies. The most famous example of this are the Dutch and English East India companies, that held a monopoly over the import and sale of Indian and Indonesian products on the European markets, and that were responsible for the first Dutch and English expansion in the East.\footnote{These early examples of joint stock companies were given by their governments a special charter, granting them a monopoly over trade with the East. This meant that no Dutch or Englishmen could legally compete with the national companies for this lucrative trade. This strategy to put the trade monopoly in the hands of private individuals stood in contrast with the Portuguese tradition, in which the monopoly had remained mostly in the hands of the Crown (see Disney, 2009, pp. 149-153), and proved to be much more successful. Of course, the charters only shielded the companies from competition with national traders, but not from competition with chartered companies or traders from other countries.}

In some other cases, the monopoly was a “national” monopoly, open to all citizens of the mother country. Examples of this include the laws that regulated the trade of English and Spanish America in the 17th-19th century. By the English Navigation laws, goods imported by the American colonies had to come from Britain, and only imperial vessels and merchants were allowed to carry out trade with the American colonies. At the same time, most American exports had to be directed to Britain first, from where they could then be re-exported to other European destinations. The Spanish monopoly worked in a similar way. From the late 16th century to the late 18th century, all trade of the Spanish American colonies had to be carried out through the Spanish port of Cadiz, and Spanish manufactures had a highly preferential access to colonial markets. Similarly restrictive regulations were in force in the Portuguese and French empires in this period.

In later centuries, the capture of colonial trade by the mother country was less commonly done through outright prohibitions to trade, and more through a system of preferential tariffs. For example, French and Italian colonies in the late 19th century charged a lower tariff on imports from the mother country than from the rest of the world (Bairoch, 1989, pp. 114-116 and 122).

That imperial powers took advantage of the position to monopolize (or otherwise restrict) colonial trade is broadly in line with the result of Proposition 2. But who gained and who lost from colonial trade monopolies? From Proposition 1, we know that, had colonies been free to set their own trade policy, they would have liked to trade with all possible buyers of their exports, thus obtaining better terms of trade. Indirect evidence that this would be the optimal outcome
comes from the fact that colonists actively tried to evade the regulations imposed by the imperial monopolies.\footnote{Smuggling was huge in Spanish America, which conducted illegal trade with the Dutch, French and English through the Caribbean in the 17th and 18th century. It was also important in the colonial US, where the Northern colonies traded illegally with the French and Dutch West Indies in the 18th century.} Thus, there was an element in imperial trade regulations that transferred welfare from the colonies to the mother country.\footnote{To partially offset this transfer, colonies were often also granted privileges in the market of the mother country. Implicitly, I am assuming that the value of these privileges was small relative to the value of the privileges accorded to the mother country in the colonial market.} However imperial regulations also implied a transfer of welfare from excluded European powers to the mother country, and the guarantee that colonial trade would not fall under control of another European power. As discussed in the previous sections, these was a potential source of gain for colonists as well.

To the extent that it allowed an imperial power to effectively legislate in a colony, direct control provided a clear advantage when trying to impose a trade monopoly over the colony. But was it really a necessary pre-condition for that? After all, a European power could attempt to establish a trade monopoly by external military force. To some extent, this is what the Portuguese attempted to do in the Indian Ocean in the 16th century. Throughout the century, they managed to gain some degree of control over the spice trade by establishing a series of fortifications around the key ports (Disney, 2009, p. 146).\footnote{At the end of the sixteenth century, only five of the twenty-four significant components of the Portuguese Indian Ocean empire - Goa, Damao, Bassein, Chaul and Colombo - possessed associated territories and rural populations of any significance (Disney, 2009, p. 146).} These stronghold put the Portuguese in a better position to rebuff local rebellions, and to constrain the activity of competing local traders.\footnote{While the Portuguese monopolized the Cape route and managed to constrain the activity of competing overland traders, Findlay and O’Rourke (2007, p. 156-157) point out that throughout the 16th century, the spice trade within Asia - which the Portuguese did not get any close to controlling - remained much more important than that with Europe.} In many cases, however, competition from other traders made this difficult to accomplish, since local populations had an incentives to ally themselves with competitors to break out of the monopoly. In fact, as competition intensified in some parts of the Indian Ocean with the arrival of the Dutch and English, it became increasingly difficult for Europeans to enforce a trade monopoly without extending the degree of formal control over producers.\footnote{For example, English interference was a key factor leading up to the first Dutch territorial acquisition in the Far East - the island of Neyra in the Banda archipelago (1609; see Masselman, 1963, pp. 260-263).}
Atlantic Europe from the 16th century onwards, and that centralized empires can be interpreted as institutions through which competing European powers attempted to appropriate part of this trade for themselves. But how were the phases of imperial expansion determined by the interactions of the European powers, and how did this depend on economic fundamentals? I address these questions in the next three sections.

4 - Phase I equilibrium

In Proposition 3, I found that intense competition for trading with $C$ may make imperial control very profitable for $B$ and $F$, creating an equilibrium in which the two countries stage war on each other, and $C$ is always subjugated. This result may be useful to understand the first phase of European imperial expansion, 1492-1770.

The countries of Atlantic Europe entered the era of discoveries fragmented in a handful of independent kingdoms,\(^{37}\) often in competition with one another for supremacy in Europe. Initially the only two countries of Atlantic Europe with enough skills at oceanic navigation, Portugal and Spain found a way to reduce competition among themselves by having the Pope\(^{38}\) attribute each of them exclusive territorial rights over half of the New World: to Spain the Americas and the very Far East (mostly the Philippines), to Portugal Brazil, the coast of Africa, and the whole of the Indian Ocean.

While territorial expansion in the East was very limited in this period - partly because of the de facto monopoly that the Portuguese enjoyed on the Cape route, partly because of the relative strength of local populations - territorial expansion in the West was swift. As the Spanish and Portuguese economies declined relative to England, France and the Netherlands in the 16th and 17th century, imperial control of the Americas would prove very profitable for these two countries, since it allowed them to monopolize a colonial trade that would have otherwise been largely foreign.\(^{39}\)

\(^{37}\)This was particularly true after the Peace of Westphalia of 1648, which sanctioned the end of the medieval, universalistic aspirations of the Holy Roman Empire and the beginning of an era of European nation states (Kissinger, 1994, pp. 56-78). In the same year, the end of the 80-year war brought the Netherlands full independence from Spain.

\(^{38}\)This happened in the famous Treaty of Tordesillas of 1494.

\(^{39}\)By imperial regulations, all Spanish America’s trade with to Europe had to pass through the Andalusian port of Cadiz. By the late 17th century, however, just about 5% of the goods leaving Andalusia were of Spanish origins,
The 17th and 18th century marked a significant increase in European competition for Empire, as the Netherlands, England and France - who did not recognize the Papal decree of 1494 - entered into the picture. These three countries had the most dynamic economies of Europe in this period (Wallerstein, 1980), and, with important fluctuations, were at a comparable level of development in terms of manufacturing and shipping capacity (see Figures 11 and 12). As a consequence, they competed with one another for control of colonial trade. In the East, they competed mainly to intermediate exotic commodities which were paid for in silver, while the market for European manufactures was very thin.\textsuperscript{40} Commodities such as the spices, coffee and tea, had a large market in Europe, and countries who could monopolize their imports stood to gain a lot from their intermediation. In the West, they competed both to intermediate exotic commodities with a European-wide demand (chiefly silver, sugar and tobacco) and to export manufactures.\textsuperscript{41}

![Urbanization rates](image1)

![Per capita GDP](image2)

![P.c. level of industrialization](image3)

Figure 11: \textbf{Countries of Atlantic Europe, relative development in 16th-18th century}

Intense competition for colonial trade resulted in frequent wars among the imperial powers of Atlantic Europe, in particular between the Dutch and the English (in the 17th century) and the Andalusian merchants ... “had been turned into nothing more than the agents of foreign manufacturers and businessmen” (Walker, 1979, p. 11 and 13). Thus, imperial control greatly benefited the Spanish merchants in this period, as well as the Spanish government (who duly taxed the transiting colonial trade).\textsuperscript{40}

The fact that colonial imports were paid for mainly with specie resulted in a strong balance of trade deficit of England vis-a-vis the East.\textsuperscript{40} The latter element was naturally more important in colonies of settlements such as British North America and, to a lesser extent, French North America and Spanish South America. Because of the importance of trade in manufactures, English trade with America was much more balanced than that with East India.\textsuperscript{41}
between the English and the French (18th century). One of the key goals of these wars was to establish, maintain or expand imperial control over as many overseas territories as possible. In turn, this led to the engrossment of national profits from colonial trade by the establishment of national monopolies of the trade over the conquered territories. This was worth fighting a war for policy makers at that time. On one hand, control of colonial trade gave European governments the financial resources they needed to consolidate their military power, in a period when national fiscal systems were not yet capable to do so (Howard, 1976, p. 38). On the other, control of colonial trade created huge profit opportunities for domestic groups, chief among which the merchants. Not surprisingly, these were given a key role in the wars of the 17th and early 18th century - being often the ones in charge of engaging the enemy in the high seas\textsuperscript{42} - and where often \textit{against} the termination of hostilities.\textsuperscript{43} More in general, the conquest of colonial trade

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure12.png}
\caption{\textbf{England and the Netherlands, relative commercial development in the 16th-18th century.}}\label{fig:enandnl}
\end{figure}

\begin{table}
\centering
\begin{tabular}{|c|c|c|}
\hline
Year & England & the Netherlands \\
\hline
1567 & 250 & 50 \\
1686 & 1000 & 750 \\
1780 & 2000 & 1500 \\
\hline
\end{tabular}
\caption{Shipping tonnage ('000 tons) and population of the three largest ports ('000).}
\end{table}

\footnotesize\textsuperscript{42}For example, the basic understanding underpinning the concession of monopoly right to chartered merchant companies was that the state would appropriate part of their profit, and that the companies would act as naval subcontractors against rival powers (Ferguson, 2011, p. 38). This military role played by the merchants was partly due to technological reason: gun power in the 17th and early 18th century was diffused just enough to make a well equipped merchant ship almost equivalent to a warship. As the intensity of gun power became more and more important in the 18th century, the power gap between the two type of ships widened again, leading to a substantial professionalization of the navy (Howard, 1976, p. 41).

\footnotesize\textsuperscript{43}For example, for the Dutch merchants “the twelve-year truce with Spain between 1609 and 1621 was an unwelcome interruption to a course of unbroken predatory success”, and “when in 1640 Portugal again separated itself from the Spanish crown and sued for peace, both the East and the West India Companies petitioned against it”. Similarly, “we find British merchants in 1745 viewing the prospect of peace with France and Spain with quite
was seen as part of a virtuous circle, whereby controlling more trade made a country richer and more powerful, which in turn would allow it to control more trade. Not surprisingly, the doctrine became endemic throughout Europe according to which governments should be “...concerned to build closed, antagonistic, mutually exclusive trading systems, each of which could flourish [...] only on the ruins of the other” (Ibid., p. 46).

Thus, we can use Proposition 3 to hypothesize that the great expansion of centralized colonial empires in the 16th-18th century - as well as the troubled state of international relations in that period - were in part determined by the political and economic fundamentals that created high European competition for colonial trade. But how would have a counterfactual situation played out? With many caveats, we can address this question by looking at the alternative equilibrium that realized in Phase II of European imperialism, 1830-1880. This is the goal of the next section.

5 - Phase II equilibrium

After the serious setbacks of the age of the American Revolutions (1776-1827)\textsuperscript{44}, European expansion continued in 1827-1880, but in a very different form than in the first era of expansion. Four big differences stood out. First, the industrial revolution transformed Britain into a clear economic leader. Second, Britain unilaterally adopted free trade in the first half of this period, triggering reciprocal trade liberalization in Europe and in the colonies. Third, Continental Europe’s imperial expansion was very limited, while British expansion became very different in nature, preferring techniques of “informal” domination to explicit administrative control. This “new” type of imperialism has been famously named the “imperialism of free trade” by Gallagher and Robinson (1953). Finally, the British Royal Navy stood up as a clear military leader in this period, contributing to ensure stability in international relations. In this section, I will see how these facts can be related to one another using the mechanism described in the model.

\textsuperscript{43} as much alarm as the Dutch had viewed the prospect of peace with Spain and Portugal a hundred year earlier” (Ibid., p. 43, 48).

\textsuperscript{44} Following to revolution in the American possessions of Britain (1776), France (Haiti; 1791), Spain (1810-1827) and Portugal (1821), Atlantic Europe’s possessions in the Western Hemisphere were reduced to Canada, most of the Caribbean, and very small parts of South America. See Bonfatti (2011) for a study of the economic determinants of rebellion in British and Spanish America.
The mass mechanization of production in Britain from the 1790s onwards brought sweeping changes to international trade patterns. As British manufactures supply boomed in 1790-1860, the price of these goods collapsed, and British exports boomed in this period. This expansion was rather evenly spread across world and European markets, resulting in a strengthening of the pattern of international specialization. During this period, most British trade partners further specialized in the production of food and raw materials, and several countries in the European periphery underwent a period of rapid de-industrialization (Williamson, 2008). As a result of international specialization, the share of British manufactures in world markets increased substantially.

![Figure 13: Share in world manufacturing, selected countries 1800-1953. Source: Bairoch (1982).](image)

Figure 13: **Share in world manufacturing, selected countries 1800-1953.**
Source: Bairoch (1982).

Industrial development tilted British trade policy in favor of the industrial sector. As consumers of foodstuff and raw materials, these sectors primarily wanted a decrease in the import tariff on these goods. This clashed with the interests of the landed elite, to whom a high import tariff on wheat guaranteed large agricultural rents. More in general, the industrialists perceived the dismantling of domestic and imperial protection as a way to encourage other European countries to follow suit, thus opening up the European markets for British manufactures. As the industrialists became more influential, several steps towards trade liberalization in Britain were taken in the 1820s and 1830s;\(^\text{45}\) as for the colonies, the traditional system of quasi-prohibition

\(^{45}\)For example, custom duties on raw materials and industrial products were reduced in 1822, and import
to trade with foreign countries was replaced with a *preferential* system (1822-1825), whereby the colonies erected discriminatory (but not prohibitory) tariffs on their import of manufactures from foreign countries and Britain did the same with the import of foodstuff and raw materials.\(^{46}\) Then, in the 1840s trade liberalization experienced a sharp acceleration. In Britain, the contentious import tariff on wheat was substantially decreased in 1846. In the colonies, the preferential system was eliminated in 1849, marking the transition to an open door system (Bairoch, 1988, p. 110). These reforms inaugurated the era of British free trade, which would last essentially untouched until World War I.

Britain’s unilateral steps towards liberalization were initially not matched by reciprocation in Europe, where manufacturing interests successfully lobbied for protection against British competition. By mid-century, however, the acceleration of liberalization in Britain combined with the rising importance of European export industries shifted the balance towards reciprocation. There was a period of about 20 years (from the late 1850s to the late 1870s) in which, supported by a series of bilateral treaties (mostly between Britain and other countries), free trade triumphed in Europe. This lasted until the manufacturing centers of the continent did not begin to catch up with Britain, and won new protectionist measures from the 1880s onwards.

As British economic power expanded and British-led trade liberalization took hold, a vocal opposition to direct imperial rule began to emerge in Britain. The same people who advocated free trade - by and large people with connection to the manufacturing sector - were often against direct imperial rule, which they portrayed as an expensive outlet for the British aristocracy. This did not mean that these people rejected the concept of British “imperial” domination over the world. Instead, they sought that this goal could best be achieved through overwhelming economic and political influence, rather than through direct administrative control (Semmel, 1970).\(^{47}\) By this, it was meant that British imperial policy should focus on keeping foreign markets *open* for British business (as opposed to *preferentially open*). This did require the use

---

\(^{46}\) For example, colonial import duties on foreign woolen manufactures dropped from 50% to 15%; on glass from 80% to 20% per cent; and on iron £6 10s. to £1 10s. per ton. The maximum tariff for foreign manufactured articles was fixed at 30% (Bairoch, 1988, p. 109).

\(^{47}\) For example, “One Whig, speaking before the House of Commons during the Corn Law debate of 1846, described free trade as the beneficent ‘principle’ by which ‘foreign nations would become valuable Colonies to us, without imposing on us the responsibility of governing them’” (Semmel, 1970, p. 8).
of economic and military leverage to induce independent local rulers to an open door policy, and to respect British investments; and it also required to maintain a clear naval supremacy over competing European countries, so as to prevent foreign annexation of overseas territories. However this new imperial policy, differently than earlier ones, did not necessarily imply the extension of formal administrative rule. This was, perhaps, because there were in every foreign country influential export groups that stood to gain from cooperating with Britain at maintaining a free trading system. In summary, “... this willingness to limit the use of paramount power to establishing security for trade is the distinctive feature of the British imperialism of free trade in the 19th century, in contrast to the mercantilist use of power to obtain commercial supremacy and monopoly through political possessions” (Gallagher and Robinson, 1953, p. 6).

In this era of imperialism of free trade, the expansion of direct imperial rule was avoided when possible. At the same time, almost all British settlers colonies (Canada, Australia and New Zealand) were granted responsible government, a form of administrative autonomy that implied a significant degree of independence for the colonies.48 As a result, the area of the directly administered British empire increased very slowly in this period, as compared to earlier or later periods. At the same time, the overwhelming power of the Royal Navy contributed to preventing other European countries from expanding as well. This pattern is illustrated in Figure 1. Particularly because of a rapid expansion of the frontier in Canada and Australia, the size of the formal British Empire increased dramatically in 1830-1880. However if we look at the directly administered empire (that is, excluding the more autonomous colonies) this increase is much more modest. Figure 1 shows that, at the same time, the residual empires of Atlantic Europe - which were drastically reduced between 1760 and 1830 by the independence of Spanish and Portuguese America - did not grow or grew only marginally. Nor did the “new” European and non-European imperial powers obtain any territorial gain in this period.

My model provides the following interpretation of the facts described above. Because of the industrial revolution, the British market share as a seller of manufactures and buyer of raw materials increased almost everywhere, creating strong international complementarities that led

48In particular, Bairoch (1988, p. 110) argues that, from the 1860s onwards, these countries “should be considered independent as far as trade policies are concerned”.

46
to the liberalization of trade. Naturally, this open door policy extended to both the domestic market and the markets of the British Empire. As British colonial trade policy became one of open doors, the trade-related need for direct administrative control of colonial or foreign territories weakened. However a weakened need for direct administrative control did not mean that Britain could simply withdraw from Empire. Overseas markets could still be lost to British trade, either because independent overseas rulers may wish to implement import substitution policies and because other European powers may wish to take over the weakest among the overseas territories. Still, these two risks could more easily be dealt with through a policy of overwhelming military power, and the extension of informal influence over overseas trade partners.

It is interesting to notice that, while British naval supremacy was by and large successful at preventing the overseas expansion of other European countries, this was perhaps also due to the political consolidation taking place in Europe at that time. On one hand, industrial followers like Germany only became a single political entity during this period. On the other, the re-unification of Germany and Italy was in itself a substitute for the expansion of markets through colonial conquest. Alesina, Spolaore and Wacziarg (2000) argue that protectionism in the first half of the 19th century created strong economic incentives for the various German and Italian regions to unify. In my argument, it was the need for protection that called for national re-unifications, as import tariffs were likely to be more effective in national markets than in regional ones.

**6 - Phase III equilibrium**

After the 1870s, a new era of European overseas expansion rapidly came about (see Figure 1). Of the old imperial powers of Atlantic Europe, only Britain and France participated in this overseas expansion, while Portugal, Spain and the Netherlands remained on the sidelines. In their stead, Belgium, Germany, Italy, and, outside Europe, Japan and the US, entered the imperial race. Following to rapid expansion in 1880-1914 and to the peace terms of World War I, the size of colonial empires reached a new zenith by the 1920s. In particular, old and new European imperial powers expanded to control the whole of Africa and large parts of Indochina. One striking fact about this expansion was that Britain herself - that for decades had been reluctant
to expand the size of her formal empire, and so remained even in the first stages of the expansion - ended up having the better part of the new territories (Cain and Hopkins, 2001).

The imperial expansion of the late 19th century was associated with very different economic conditions relative to the previous period. Most notably, the new imperialism began sometimes after the industrial revolution had spread to various countries outside Britain. In fact, one thing that all the new imperial powers and France had in common, was that they were all rapidly catching up with Britain in terms of industrial development. As a consequence of this catching up, British industrial leadership - which had reached its zenith in the 1860s - began to decline in the 1870s, eventually being replaced by the United States in early 20th century. The emergence of several industrial powerhouses in Europe and elsewhere meant that the level of competition for overseas markets - both for exports and for imports - intensified markedly. For Britain, this meant the beginning of a century-long period of decline in her market share.

The model suggests one way to relate the new imperialism of 1880-1914 to the decline in the British share of world markets beginning in the 1870s. An enhanced level of competition for the markets of non-industrial countries had two consequences for the incentives of industrial countries to expand, or create, an empire. First, the expansion of international trade as a share of national income may have scaled up the attractiveness for all industrial countries of securing some overseas markets through military conquest (at least to the extent that it grew faster than the cost of armament, as a share of national income). Second, Britain’s loss of market supremacy removed the stabilizing factor that had characterized international relations in the previous period.

The argument that enhanced competition for overseas markets can contribute to explain 19th century imperial expansion is, of course, not new. In fact, there has been a fierce debate among historians over whether the late 19th century expansion was due to economic or other factors, particularly for the case of Britain.49 I abstract from such a debate, since I am in no position to

49Two main alternative explanation for British expansion in Africa have been put forward. First, British expansion was mostly strategic, in that it aimed at protecting the eastern flank of the Indian Ocean from foreign intrusion (the famous Cairo to Cape axis of British strategy) or at avoiding an excessive expansion by other European powers (West Africa). Second, a series of political crises in African countries (such Egypt’s bankruptcy in the late 1870s) disrupted British interests in these countries to the a point that the British government opted for a direct takeover. The case for a non-economic interpretations was forcefully made by Gallagher and Robinson (1961), who were then followed by a series of works in a similar mood, see e.g. Fieldhouse (1973). In recent years,
prove either side wrong. My main goal in this paper is to incorporate some of the main intuitions of the economic argument - together with insights from the international relation literature - into a formal model. This has two main benefits. First, it clarifies the assumptions made, and provides guarantees that the argument is internally consistent with these assumptions. Second, it may lead to the generation of finer empirical predictions. Ultimately, a finer understanding of all the arguments put forward, together a more theory-driven empirical work, is likely to be the way forward for adjudicating the debate on the origins of late 19th century imperialism.

7 - Empire and globalization

The gist of the Corollary to Proposition 2 is that having constructed a larger “domestic” market for her exports, a successful imperial power has a stronger incentive to restrict its own imports of competing products. This result may contribute to explaining the fact that two key eras of imperial expansion (the 17th-18th century, 1880-1914) were associated with a deterioration in the trade relations between European countries.

Until the late 17th century, there had been an active intra-manufacturing trade between the main industrial centers of Europe. For example, as late as 1699-1701, imports (exports) of manufactures to (from) Northwestern Europe accounted for 17% per cent (32%) of total British imports (exports; see Figure 14). In the first half of the 18th century, however, a strong movement towards industrial self-sufficiency took place in Europe, supported by a widespread rise of import tariffs (Davis, 1966). As a result, intra-manufacturing trade in Europe collapsed: in England, for example, imports (exports) of manufactures from (to) Northwestern Europe declined even in absolute value in 1699-1774 (see Figure 14), despite a rapid increase in the trade and income of the countries involved.

---

50 For example, the claim that imperial expansion should be causally associated with a rise in protectionism - see the next section - is, to the best of my knowledge, new to the literature. If anything, the earlier literature had claimed the opposite, that is that the rise of protectionism made the European countries look markets elsewhere, and this led to imperial expansion. For a criticism of this view, see Fieldhouse (1973), Ch. 2.

51 This intra-manufacturing trade was dominated by the export of woolens from England and of that of linens from the Netherlands and France. For the data, see Davies, 1962, pp. 300-303.

52 According the Maddison (2006), GDP increased by 82% in France in 1700-1820, 238% in Britain, 6% in the Netherlands, and 96% in Germany.
Although the original impetus behind the tariff rise was either fiscal or political - major intra-European wars took place in this period, stretching the need for government revenues and demanding trade retaliation against European enemies - the successful development of protected industries soon created vested interest that lobbied for the continuation, or even the reinforcement, of what eventually became protective tariffs (Davis, 1966). These industries were made particularly successful by the expanding size of the “domestic” market, which extended beyond the home market to include rapidly growing colonial empires. In line with the model’s result, colonial protection created a force that pushed for more protection against European manufactures in Europe, and not just in the colonies.

Figure 14: English manufacturing trade with NW Europe (Germany, Holland, Flanders and France), 1666-1773.
Unit: absolute values are in million £. Source: Davis (1954, 1922).

Similarly, the late 19th century imperial expansion was shortly followed by a rise of protectionism in the main imperial powers. As in the mercantilistic era, this was based on the two-pronged strategy to simultaneously restrict extra-imperial trade in the metropole and the empire, while keeping intra-imperial trade relatively free. Trade restrictions increased constantly in the run-up to World War I, and in the 1920s. With the coming of the Great Depression, this process accelerated markedly. In Britain, where free trade had been the mainstay of policy for almost half a century, imperial expansion was not associated with a return to protectionism in 1880-1914, although domestic pressure in that sense was clearly increasing. It was only after the WWI that the first protectionist policies were adopted. In the 1930s, Britain ended up restricting
extra-imperial trade much more than many other countries.

The model’s results would suggest that part of this increased protection was due to the creation of larger “domestic” markets for imperial manufactures. Of course, there were other factors, unrelated to the rise of the new imperialism, that played a key role in determining the rise of protectionism - chief among all, the Great Depression of the 1930s. Notice, however, that this is not necessarily in conflict with my interpretation. Whatever shock a country experiences that calls for an increase in protection, this is likely to have a stronger effect when domestic producers have a recently expanded internal market at their disposal, and when the government cares less about consumers in this market than it did in the past. In other words, the Depression may well have had such a strong hit on international trade because some of the key players in the international trading system seated at the heart of large colonial empires.

5 The end of European imperialism, 1945-1990

Following to World War II, colonial empires that had existed for centuries collapsed within less than two decades. Immediately after the war, and despite the difficulties of reconstruction, countries like Britain actively tried to preserve their empires and the preferential trading systems that had governed them (again) since at least the 1930s. By the mid 1950s, however, it appeared clear that the most dynamic centers of growth for European trade were in Europe, the US, and Japan and not in preferential colonial markets (Cain and Hopkins, 2001, p. 632). Decolonization followed shortly thereafter, in most cases with no opposition on the side of imperial powers.

Figure 10 illustrates that the share of British exports to European countries plus the US and Japan increased exponentially after 1945, reaching 90% of British exports by 1985. This reduced the importance of trade with the less developed rest of the world - that rest of the world over which Britain and other European powers had fought for well over four centuries - to the same order of magnitude as in mid 17th century.

But what economic forces were responsible for the rapid increase in the relative importance

---

53In the case of Britain, the attempt to keep the empire together was made even more difficult by the loss of India. Still, Cain and Hopkins (2001, p. 630) argue that this loss did not, per se, provide a sufficient motivation for the rest of the British empire.

54The case of Algeria and the Portuguese colonies being the obvious exceptions.
of trade between industrialized countries? One key change in the structure of international trade after World War II was triggered by a dramatic rise in the importance of *intra-industry* trade in manufacturing (Grubel and Lloyd, 1975). This arose following to technological change in existing industries, and to the birth of many new “modern” industries. A key feature of the new technologies is the existence of increasing returns to scale, that make firms the sole producers of the specific variety of product that they design through initial investment. This regime of monopolistic competition between firms ensures that goods exist in an increasing number of varieties, and that these may be located in many different countries at the same time. This leads to an increase in two-way trade between countries within each specific manufacturing industry, and to an increase in total trade between industrial countries.

This pattern is illustrated in figure 15, which plots the share of two sets of industrial countries in UK trade in 1948-2010. The figure shows that the share of these countries in UK imports and exports increased much faster than their share in world GDP, pointing to a structural change in the nature of trade between industrialized countries.\(^{55}\)

![UK share of trade with other industrial countries, 1948-2010.](image)

**Figure 15:** UK share of trade with other industrial countries, 1948-2010.
Sources: trade, IMF DOTS; GDP, Maddison (2001).

\(^{55}\) Of course, faster trade liberalization among the advanced economies is another factor explaining this pattern, as may be decolonization itself. In any case, a very rapid increase (and from relatively levels) in intra-industry trade after WWII is well documented in the trade literature (see Grubel and Lloyd, 1975). The only work on pre-WWII intra industry trade that I am aware of (Jacks, 2011), finds that Canadian intra industry trade remained constant and relatively low in between the two wars.
The results derived under the assumption that $\sigma$ is low (see section 3) may then be useful to interpret the economic forces that led to decolonization. As manufactured goods became increasingly more differentiated after the War, trade with other manufacture producer became the main source of gains from trade for European imperial powers. This weakened the rationale for defending preferential trade agreements with raw-material exporting colonies, particularly since these preferential trade agreements could have the undesirable effect to disrupt the trade relations of the mother country with other Western countries. Interestingly, the model suggests that, in this case, the existence of an industrial leader - say the US - makes the followers even less interested in defending their empire, since the presence of a leader now creates stronger trade opportunities outside of the empire. In this interpretation, decolonization happened because the old terms-of-trade motivations for having a colonial empires were made obsolete by changes in technology and trade patterns. And the collapse of colonial empires - for centuries the mainstay of European mercantilism - may in itself help to understand the great opening to international trade between industrial nations that was achieved with the European Union and the GATT/WTO.

6 Conclusions

The history of European imperialism is a fascinating topic, and one of crucial importance for modern economists. European imperialism shaped the world economy for almost five centuries, leaving a legacy that is known to matter for current development experiences. In this paper, I have sought to improve our understanding of the economic forces that led to the rise and decline of European imperialism, and of the impact of imperialism on contemporary economic conditions. I have built a simple model of the benefit of imperialism, which describes a world in which colonies exchange raw materials for manufactures with the rest of the world, and what defines them as colonies is the fact that their trade policy is controlled by an external imperial power, itself a manufacture exporter. Having determined how the imperial power’s gain from empire depends on the economic fundamentals that shape the pattern of trade, I have then divided the history of European imperialism into four main phases, and used my results to argue that these may be accounted for with fluctuations in the benefit of imperialism. I have also noticed that phases
of imperial expansion were associated with periods of restricted trade between the European countries, while phases of imperial relaxation or contraction were associated with periods of globalization. The model offers a causal interpretation for this link, by showing that some degree of trade restrictions between European countries must be put in place if wealth is to be extracted from the colonies.

In a nutshell, my main results are as follows. The discovery of the Americas and of the sea way to the Indies created huge new trade opportunities for Europe. Centrally administered colonial empires were one institutional arrangements through which a few European countries envisaged to secure a disproportionate value of that trade for themselves. Fluctuations in the value of such institutions may contribute to explaining fluctuations in the size of colonial empires in the 16th-20th century. In particular, imperial institutions had a particularly high value in periods in which, because of strong economic competition between the imperial powers, colonial trade was highly contested - for example, in the two eras of great imperial expansion (16th-18th century; and 1880-1914). On the contrary, when colonial trade was dominated by an industrial leader - such as was Britain during the industrial revolution - the return to imperial institutions declined for this country. This, together with the leader’s overwhelming military power, may help explaining the relaxation of imperialism in 1820-1880. Finally, the rise of diversified manufactures and intra-manufacturing trade in the 20th century marked a structural break in this pattern, by decreasing the value of colonial trade relative to trade among the imperial powers. This decreased the return to imperial institutions for all levels of competition, possibly explaining their final abandonment.

The model presented in this paper is very simple, and yet the historical phenomena that it seeks to explain were vast both in space and in time. It follows that there are many interesting questions that have not been considered here, and one could think of a range of extensions to begin to address them. I will here briefly mention three.

First, my focus in this paper was on a specific set of imperial institutions, that is centrally administered ones. Implicitly, I assumed that direct control is required to enforce a local preference for the imperial power’s own economy, which goes to the detriment of both local exporters and international competitors. There may, however, be other trade-related goals that can be pursued
through external pressure, and the end of direct control may then lead way to a more informal set of imperial institutions. For example, external pressure - perhaps with the complicity of local export interests - may be helpful to prevent local governments from adopting import-substitution policies, or to disrupt foreign investments in the export sector of the economy. In fact, there is historical evidence that informal imperial techniques were a feature of British mid-19th century imperialism, and of Western foreign policy in third world countries during the Cold War. One natural extension of the current model would be to add informal empire as a third possible set of institutions, and see how the emergence of these is determined by the political economy of policy formation in C.

Next, the assumption of discrete trade policy restricts the scope of the paper in two important ways. On one hand, the paper provides only very stylized indications as to the relation between imperial expansion/contraction and protectionism. Instead, to introduce continuum trade policy could generate testable predictions for the impact of specific imperial evolutions on the trade policy of the imperial power in specific industries. Since good data on trade and trade policy exists from 1870 onwards, one could use these predictions to estimate the impact that late-19th century imperialism, or decolonization, had on European trade policy. On the other hand, the assumption of discrete trade policy does not leave any room for modeling trade talks between the two imperial powers. However, very important rounds of trade talks took place around the crucial times of imperial relaxation, that is in mid-19th century and after World War II. Furthermore, some of the arguments made in this paper do point to a link between the existence of colonial empires and the outcome (and occurrence) of trade talks - if anything, the suggestion that trade restrictions are a consequence of imperial expansion is clearly relevant. To include these issues could help shed light on the historical factors that led to the creation of GATT/WTO, the multilateral institution that underpins the current era of globalization.

Finally, one important fact that the current model is not considering is that the construction of colonial empires was organized very differently across space and time. In some cases, the state outsourced the imperial enterprise to specific stake-holders - merchants, investors, etc - by rewarding them with colonial assets and with monopoly positions in colonial trade. In some others, the state was directly involved in establishing and maintaining imperial control. In these
latter cases, the national monopoly over colonial trade was normally a “public monopoly” to which all metropolitan citizens could, in principle, participate. While there seems to be a broad pattern as to the evolution of the organization of imperialism over the centuries - from private to public - there are interesting exceptions to this, such as the occupation of some African countries by chartered companies in the late 19th century. The model could be extended to allow for two forms of empire-building - public and private - which are then associated with different regimes of colonial trade. One could then investigate what is the welfare maximizing arrangement, and how does this depend on trade conditions. By determining which groups in the European society are poised to gain more from the imperial enterprise, this could shed further light on the relation between colonial trade and the evolution of political power in Europe.

Appendices

A Derivations

Uncompensated demand

Setting \( \text{MRTS}_{a,b} = \left( \frac{y_d}{y_d} \right)^{\frac{1}{\sigma}} = \frac{y_d}{y_d} \) and plugging in the budget constraint yields \( y_a^I p_a^I + \left( \frac{y_d^I}{y_d^I} \right)^{\sigma} y_a^I p_b^I = (1 - \mu) m^I \). Re-arranging, we obtain:

\[
\begin{align*}
y_a^I &= \frac{(1 - \mu)m^I}{p_a^I (p_a^{1-\sigma} + p_b^{1-\sigma})} \\
y_b^I &= \frac{(1 - \mu)m^I}{p_b^I (p_a^{1-\sigma} + p_b^{1-\sigma})}
\end{align*}
\]

Equilibrium prices

Suppose that the set \( M \) of countries trade. Then, for the \( y_a \) and \( y_a b \) market to be in equilib-
rium we must have:

\[(1 - \mu) m^M \left( \frac{p_a^M}{(P^M)^{1-\sigma}} \right)^{-\sigma} = y_a^M \quad (13)\]
\[(1 - \mu) m^M \left( \frac{p_b^M}{(P^M)^{1-\sigma}} \right)^{-\sigma} = y_b^M \quad (14)\]

Dividing (13) by (14), we obtain \(p_b^M = \left( \frac{y_a^M}{y_b^M} \right)^{\frac{1}{\sigma}} p_a^M\); plugging this in the condition for equilibrium in the \(x\) market we obtain:

\[
\mu \left( x^M + y_a^M p_a^M + y_b^M \left( \frac{y_a^M}{y_b^M} \right)^{\frac{1}{\sigma}} p_a^M \right) = x^M
\]

Re-arranging (and proceeding identically for \(p_b^M\)), we obtain:

\[
p_a^M = \frac{1 - \mu}{\mu} \frac{x^M}{(y_a^M)^{\frac{1}{\sigma}} \left( (y_a^M)^{\frac{\sigma-1}{\sigma}} + (y_b^M)^{\frac{\sigma-1}{\sigma}} \right)}
\]
\[
p_b^M = \frac{1 - \mu}{\mu} \frac{x^M}{(y_b^M)^{\frac{1}{\sigma}} \left( (y_a^M)^{\frac{\sigma-1}{\sigma}} + (y_b^M)^{\frac{\sigma-1}{\sigma}} \right)}
\]

**Indirect utility**

Plugging uncompensated demand into direct utility gives:

\[
v = \left( \frac{m}{2} \right)^{\frac{1}{\sigma}} \left[ \left( \frac{1}{2p_1^{\sigma}p^{1-\sigma}} \right)^{\frac{\sigma-1}{\sigma}} + \left( \frac{1}{2p_2^{\sigma}p^{1-\sigma}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{2(\sigma-1)}}
\]
\[
= \frac{m}{2p_1^{1-\sigma} + p_2^{1-\sigma})^{\frac{\sigma}{2(\sigma-1)}}}
\]
\[
= \frac{m}{2p_1^{1-\sigma} P^2}
\]
\[
= \frac{m}{2p_1^{1-\sigma}}
\]

**B Proofs**

**Proof of Proposition 1**
I begin by showing that within any pair of countries that trade with each other, there is always at least one country that is strictly better off if the third country is also admitted to. The strategy for the proof is to show that there always exists a lump sum transfer $T$ from $B$ to $C$ (or from $C$ to $B$), given which the three-country trading system is strictly Pareto-superior to the two-country trading system for $B$ and $C$ (provided that the two systems do not yield the same vector of prices). This is sufficient to conclude that the country that pays the lump-sum transfer is strictly better off in the tree-country trading system than in the two-country trading system. The proof follows closely Dixit and Norman (1980, pp. 76-78). Call $v^J(BC)$ the utility of country $J \in \{B, C\}$ in the two-country trading system and $v^J(W)$ its utility in the three country trading system. Call $T^J$ the income transfer that $J$ must receive to achieve utility $v^J(BC)$ at prices $p^W$ (this may be positive or negative). If we can show that $T^B + T^C < 0$, our result follows.

Country $J$ obtains the same utility when importing $m^J(W)$ and $m^J(BC)$, but the former is optimum at prices $p^W$. It follows that $T^J = p^W \cdot m^J(W) < p^W \cdot m^J(BC)$ (the strict inequality being due to substitutability among goods, and the fact that have assumed $p^W \neq p^{BC}$). It follows that $T^B + T^C = p^W \cdot [m^B(W) + m^C(W)] < p^W \cdot [m^B(BC) + m^C(BC)] = 0$, where the latter equality comes from the condition for product market clearing in the two-country trading system. I now show that $\{C, B, F\}$ is the only CPNE. There exists only two other types of trade outcomes, the first in which no country trades ($\{\emptyset\}$), and the second in which one country does not trade. Since moving from autarchy to free trade is always strictly welfare-increasing for a country (provided that it provides a different set of prices, and there is some substitutability in consumption; see Dixit and Norman, 1980, pp. 76-78) the first outcome is strictly Pareto-inferior to $\{C, B, F\}$. It follows that no-coalition can form to deviate from $\{C, B, F\}$ to $\{\emptyset\}$. As for deviating to the second outcome, this is strictly welfare-decreasing for the countries that falls into autarchy. However by our proof above it is also strictly welfare-decreasing for at least one of the two countries that continue to trade. It follows that there exist no coalition that is willing to undertake a deviation towards such outcome, and $\{C, B, F\}$ is a CPNE. To see that it is the only CPNE, notice that $\{C, B\}$, $\{C, F\}$, and $\{B, F\}$ cannot be one, since a coalition would always form between one of the two country who trades and the excluded country, to admit the latter into trade. But $\{\emptyset\}$ is not one either, since any pair of countries would be strictly better off by
forming a coalition to start trading between themselves. ■

Proof of Proposition 2

Because any trade policy opening by the imperial power (w.l.o.g., $B$) must be reciprocated by the other industrial country in a SPNE, there are three prices among which $B$ can choose: $p^{BC}$, $p^W$, $p^{BF}$. If $b > f$, $B$ is a net exporter of $y$ in all three cases. The three prices can then be ranked using the fact that $v^B(p)$ is monotonically increasing in $p$ for a net exporter of $y$. Since $\frac{b}{b+f} \geq \frac{3}{b+f}$ for $b > f$, the result follows. If $b = f$, $B$ is a net exporter of $y$ in the first two cases, has zero net trade flows in the third. Again, since $\frac{b}{b+f} \geq \frac{3}{b+f}$, the result follows. If $b < f$, $B$ is a net exporter in the first two cases, a net importer in the third. Since $\frac{b}{b+f} \geq \frac{3}{b+f}$, the relevant comparison is between $v^B(p^{BC})$ and $v^B(p^{BF})$. What we need is:

$$\frac{1 + b^2}{2\sqrt{\frac{b}{b+f}}} > \frac{1 + b^2}{2\sqrt{\frac{b}{b+f}}}$$

$$3\sqrt{b} > \left(1 + 2\frac{b}{b+f}\right)\sqrt{b+f}$$

$$2\frac{b}{b+f} - 3\sqrt{\frac{b}{b+f}} + 1 < 0$$

$$\frac{1}{2} < \sqrt{\frac{b}{b+f}} < 1$$

$$\frac{1}{4} < \frac{b}{b+f} < 1$$

Since the latter is always true for $b \in [1, 2]$ and $f \in [1, 2]$, this concludes the proof. ■

Proof of Proposition 3

First scenario. That the winner of the imperial race always subjugates the colony follows from the fact that, for $J \in \{B, F\}$, $\pi_o^J(k,k) > k^A$ for $k = 1$ (Assumption 1) and is increasing in $k$ for $k \in [1, 2]$ (Lemma 1). To complete this part of the proof, I am now going to show that the unique Nash equilibrium is always $(1, 1)$. This follows from the fact that $a^J = 1$ is a dominant strategy for each country $J$. To see this, notice that if $a^{-J} = 1$, $a^J = 1$ is best response if and only if $\frac{1}{2} [\Pi^J(k,k) - k^A] > k^M$. If $a^{-J} = 0$, $a^J = 1$ is best response if and only if $\pi_o^J(k,k) > k^A$. 59
That the latter conditions always holds has already been argued; but the former conditions always holds for similar reasons, since, \( \frac{1}{2} \left( \Pi^I(k, k) - k^A \right) > k^M \) when \( k = 1 \) (Assumption 2) and is increasing in \( k \) for \( k \in [1, 2] \) (Lemma 1).

**Second scenario.** Define \( b'' = \arg \left\{ \pi_o^B(b, 1) = k^A \right\} \) and \( b' \in (1, 2] = \min \left\{ 2, \arg \left\{ \frac{f}{b+t} \left[ \Pi^F(b, 1) - k^A \right] = k^M \right\} \right\} \). Because \( \pi_o^B(1, 1) > k^A \) (Assumption 1), \( \pi_o^B(2, 1) = 0 \), and \( \pi_o^B(b, 1) \) is continuously decreasing in \( b \) (Lemma 1), we know that \( b'' \in (1, 2) \). Because \( \frac{1}{2} \left[ \Pi^F(1, 1) - k^A \right] > k^M \) (Assumption 2), \( \Pi^F(b, 1) \) is constant in \( b \) (Lemma 1), and \( \frac{f}{b+t} \) is decreasing in \( b \), it is possible that \( b' \in (1, 2] \).

Since it is \( \pi_o^B(b, 1) > k^A \) for \( b < b'' \), \( \pi_o^B(b, 1) \leq k^A \) for \( b \geq b'' \), it follows that if \( B \) wins, she does not subjugate \( C \) iff \( b \geq b'' \). Since it is \( \pi_o^B(b, 1) > k^A \ \forall b \) (Assumption 1 and Lemma 1), it follows that if \( F \) wins, she always subjugate \( C \). It also follows that \( a^B = 1 \) is a dominant strategy for \( B \) when \( b < b'' \), while \( a^F = 1 \) is a dominant strategy for \( F \) when \( b < \min \{b', b''\} \).

To complete this part of the proof, I am now going to derive the unique Nash equilibrium for all possible values of \( b \), and show that this is consistent with the behavior described. There are four possible cases, \( b < b', b'', b' < b < b'', b'' < b < b', \) and \( b', b'' < b \). If \( b < b', b'' \), there is a (necessarily unique) equilibrium in dominant strategies, \( (1, 1) \). If \( b' < b < b'' \), \( (1, 0) \) is the unique Nash equilibrium, since \( a^B = 1 \) is a dominant strategy and \( \frac{f}{b+t} \left[ \Pi^F(b, 1) - k^A \right] < k^M \).

If \( b'' < b < b' \), there are two subcases. If \( \frac{f}{b+t} \left[ \pi_o^F(b, 1) - k^A \right] > k^M \), \( (1, 1) \) is the unique NE, since \( a^F = 1 \) is a dominant strategy and \( \frac{b}{b+t} \left[ \Pi^B(b, 1) - k^A \right] > k^M \). If \( \frac{f}{b+t} \left[ \pi_o^F(b, 1) - k^A \right] < k^M \), \( (1, 0) \) is a NE, since \( B \) is indifferent between \( a^B = 1 \) and \( a^B = 0 \) and \( F \) is, by the last inequality, also at her best response. It is also unique: to see this, notice that \( (1, 1) \) is not an equilibrium, since \( \frac{f}{b+t} \left[ \pi_o^B(b, 1) - k^A \right] < k^M \); but neither is is \( (0, 0) \), since \( \pi_o^F(b, 1) > k^A \), nor \( (0, 1) \), as \( \frac{b}{b+t} \left[ \Pi^B(b, 1) - k^A \right] > k^M \). Finally, if \( b', b'' < 0 \), \( (1, 0) \) is a NE, since \( B \) is indifferent between \( a^B = 1 \) and \( a^B = 0 \) and \( \frac{f}{b+t} \left[ \Pi^F(b, 1) - k^A \right] < k^M \). Again, it is also unique. To see this, notice that \( (1, 1) \) is not an equilibrium, since \( \frac{f}{b+t} \left[ \Pi^F(b, 1) - k^A \right] < k^M \); but neither is is \( (0, 0) \), since \( \pi_o^F(b, 1) > k^A \), nor \( (0, 1) \), as \( \frac{f}{b+t} \left[ \pi_o^F(b, 1) - k^A \right] > k^M \).

**Third scenario.** Define \( f'' = \arg \left\{ \pi_o^B(2, f) = k^A \right\} \) and \( f' = \arg \left\{ \frac{f}{b+t} \left[ \Pi^F(2, f) - k^A \right] = k^M \right\} \). If \( \arg \left\{ \frac{f}{b+t} \left[ \Pi^F(2, f) - k^A \right] = k^M \right\} \in (0, 1) \), \( f' = 1 \) otherwise. Because \( \pi_o^B(2, 1) = 0 \), \( \pi_o^B(2, 2) > k^A \) (1 and Lemma 1), and \( \pi_o^B(2, f) \) is continuously increasing in \( f \) (Lemma 1), we know that \( f'' \in (1, 2) \). Because \( \frac{1}{2} \left[ \Pi^F(2, 2) - k^A \right] > k^M \) (Assumption 2 and Lemma 1), \( \Pi^F(2, f) \) is increas-
ing in $f$ (Lemma 1), and $\frac{f}{b+f}$ is increasing in $f$, it is possible that $f' \in [1,2)$. Notice that it is $\pi_o^B(2, f) < k^A$ for $f < f''$, $\pi_o^B(2, f) \geq k^A$ for $f \geq f''$, while it is $\pi_o^F(2, f) > k^A \forall f$ (Assumption 1 and Lemma 1). It follows that $a^B = 1$ is a dominant strategy for $B$ when $f > f''$, while $a^F = 1$ is a dominant strategy for $F$ when $b < \min[b', b'']$. There are four possible cases, $f \leq f', f''$, $f' \leq f < f''$, $f'' \leq f < f'$, and $f', f'' \leq f$. If $f \leq f', f''$, $(1, 0)$ is a NE, since $B$ is indifferent between $a^B = 1$ and $a^B = 0$ and $\frac{f}{b+f} [\Pi^F(2, f) - k^A] < k^M$. Again, it is also unique. To see this, notice that $(1, 1)$ is not an equilibrium, since $\frac{f}{b+f} [\Pi^F(2, f) - k^A] < k^M$; but neither is $(0, 0)$, since $\pi_o^F(2, f) > k^A$, nor $(0, 1)$, as $\frac{b}{b+f} [\pi_o^F(2, f) - k^A] > k^M$. If $f'' \leq f < f'$, $(1, 0)$ is the unique Nash equilibrium, since $a^B = 1$ is a dominant strategy and $\frac{f}{b+f} [\Pi^F(2, f) - k^A] < k^M$. If $f' \leq f < f''$, there are two subcases. If $\frac{f}{b+f} [\pi_o^F(2, f) - k^A] > k^M$, $(1, 1)$ is the unique NE, since $a^F = 1$ is a dominant strategy and $\frac{b}{b+f} [\Pi^B(2, f) - k^A] > k^M$. If $\frac{f}{b+f} [\pi_o^F(2, f) - k^A] < k^M$, $(1, 0)$ is a NE, since $B$ is indifferent between $a^B = 1$ and $a^B = 0$ and $F$ is, by the last inequality, also at her best response. It is also unique: to see this, notice that $(1,1)$ is not an equilibrium, since $\frac{f}{b+f} [\pi_o^F(2, f) - k^A] < k^M$; but neither is is $(0,0)$, since $\pi_o^F(2, f) > k^A$, nor $(0,1)$, as $\frac{b}{b+f} [\pi_o^B(2, f) - k^A] > k^M$. Finally, if $f', f'' \leq 0$, there is a (necessarily unique) equilibrium in dominant strategies, $(1,1)$. ■

**Proof of Proposition 2bis**

Because any trade policy opening by the imperial power (w.l.o.g., $B$) must be reciprocated by the other industrial country in a SPNE, there are three trade outcomes among which $B$ can chose: $BC$, $W$ and $BF$. It is easy to see that $W$ always dominates $BF$:

$$\frac{1 + \frac{1}{b+b}}{2\sqrt{2\pi^{\frac{1}{2}}}} = \frac{2}{2\sqrt{2\pi^{\frac{1}{2}}}} \sqrt{b} < \frac{1 + \frac{3}{2b+b}}{2\sqrt{2\pi^{\frac{1}{2}}}} = \frac{5}{2\sqrt{2\sqrt{3}}\sqrt{2\sqrt{2\pi^{\frac{1}{2}}}}}
$$

As for $BC$, this dominates $W$ iff:

$$\frac{1 + \frac{3}{2b+b}}{2\sqrt{2\pi^{\frac{1}{2}}}} = \frac{3}{2\sqrt{3}} \sqrt{b} < \frac{5}{2\sqrt{2\sqrt{3}}\sqrt{2\sqrt{2\pi^{\frac{1}{2}}}}}
$$

$$\left(3 - \frac{5}{2\sqrt{3}} \sqrt{2\pi^{\frac{1}{2}}}ight) \sqrt{b} \geq 0.
$$
The term in parenthesis is clearly positive for $\sigma \to \infty$, strictly increasing in $\sigma$, and negative for $\sigma \to 1$. There then exists a $\sigma \in (1, \infty)$ such that this term is equal to zero. Notice that $BC$ dominates $W$ for all possible $b$ if $\sigma \geq \sigma$, while it is the other way around if $\sigma < \sigma$. ■

**Proof of Proposition 4**

When $\sigma < \sigma$, it is $\pi_o(k) < 0$, and no country annexes $C$ if given the option. It follows that $(1,0)$, $(0,1)$ and $(0,0)$ are all equilibria, since each country $J$ is indifferent between $a^J = 1$ and $a^J = 0$ if $a^{\neg J} = 0$, and strictly better off setting $a^J = 0$ if $a^{\neg J} = 1$. For the latter fact, it also follows that $(1,1)$ is not an equilibrium.

**C Construction of figures and descriptions**

**Figure 1**

Etemad (2007, p. 131) provides estimates of the total size of empires in 1760-1938. Since Canada, Australia and New Zealand all received responsible government from Britain in 1830-1880, while South Africa mostly received it in 1880-1913, I have excluded the former three from 1880 onwards, and South Africa from 1913 onwards. I have then assumed that the size of the Portuguese and Spanish Empires had reached their size in 1760 by 1600, while the Dutch, English and French Empires were still non-existing in 1600. I have also assumed that all Empires had shrunk to zero by 1970, except for the Portuguese which ended immediately thereafter. While these are clearly approximations (for example, the Portuguese lost some of their strongholds in the Indian Ocean in the 17th century, while contemporaneously expanding their control over Brazil; Britain and France retain some small colonial possessions up to these days; etc), they are not likely to affect the long-run pattern.

**Figure 12**

The data on shipping tonnage is taken from Ormrod (2003), p. 276, and I have interpolated the missing years. Dutch tonnage for 1636 and 1750 excludes VOC and WIC ships.

**Figure 10**

**References**


