

NBER WORKING PAPER SERIES

GLOBALIZATION AND POLITICAL STRUCTURE

Gino Gancia
Giacomo A.M. Ponzetto
Jaume Ventura

Working Paper 22046
<http://www.nber.org/papers/w22046>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue
Cambridge, MA 02138
February 2016

We thank Marta Santamaria and Jagdish Tripathy for excellent research assistance. We acknowledge financial support from the European Research Council (grant GOPG 240989), the Spanish Ministry of Economy and Competitiveness (grants RYC-2013-13838, ECO2014-54430-P and ECO2014-59805-P), the Generalitat de Catalunya (grant 2014 SGR 830) and the Institut d'Estudis Autonòmics. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Gino Gancia, Giacomo A.M. Ponzetto, and Jaume Ventura. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Globalization and Political Structure
Gino Gancia, Giacomo A.M. Ponzetto, and Jaume Ventura
NBER Working Paper No. 22046
February 2016
JEL No. D71,F15,F55,H77,O57

ABSTRACT

The first wave of globalization (1830-1914) was accompanied by a decline in the number of countries from 125 to 54. The second wave of globalization (1950-present) has led instead to an increase in the number of countries to a record high of more than 190. This paper develops a theoretical framework to study the interaction between globalization and political structure. We show that political structure adapts to expanding trade opportunities in a non-monotonic way. Borders hamper trade. In its early stages, the political response to globalization consists of removing borders by increasing country size. In its later stages, however, the political response to globalization is to remove borders by creating economic unions, and this leads to a reduction in country size.

Gino Gancia
CREI and Universitat Pompeu Fabra
Ramon Trias Fargas 25-27
08005 Barcelona
Spain
and Universitat Pompeu Fabra
ggancia@crei.cat

Giacomo A.M. Ponzetto
CREI, Universitat Pompeu Fabra,
and Barcelona GSE
C/ Ramon Trias Fargas, 25-27
08005 Barcelona
Spain
gponzetto@crei.cat

Jaume Ventura
CREI
Universitat Pompeu Fabra
Ramon Trias Fargas, 25-27
08005-Barcelona
SPAIN
and Barcelona GSE
and also NBER
jventura@crei.cat

1 INTRODUCTION

In 1820, the world was made up of 125 sovereign states. By the end of the first wave of globalization in 1914, the number of countries had fallen to 54, less than half of the initial level. The interwar period witnessed a reversal of these trends: trade collapsed and the number of sovereign states rose to 76 by 1949. Until then, political and economic integration had proceeded together. Yet, the end of World War II marks the beginning of a new era. The second wave of globalization that started after 1950 has led trade between nations to flourish to levels never seen before. But this time the process of economic integration was accompanied by different changes in the world political structure. On the one hand, the number of countries has risen to a record high of more than 190, suggesting that more trade is now accompanied by political fragmentation. On the other hand, however, there has been a proliferation and growth of international treaties and unions aimed especially at fostering economic integration, such as the WTO and the European Union.

These trends are illustrated graphically in Figure 1, which shows the historical evolution of the number of sovereign states in the world, average exports as a share of GDP and the number of members of the GATT/WTO. The data on the number of states was obtained from Butcher and Griffiths (2013).¹ However, counting the number of sovereign countries in the past poses some challenges, especially in remote and underdeveloped areas.² Hence, in Figure 2 we plot an alternative measure of political concentration: the average land size of “internationally recognized” countries and of thirteen major empires since 1830, from the Cross-National Time-Series Data Archive. The graph tells a story remarkably similar to Figure 1. During the 1800s, there was a phase of political consolidation in which countries and empires expanded their territories. But this trend reversed in the decades after World War II, which saw the collapse of large empires and a steady fall in average country size.

Why did the first wave of globalization lead to political concentration? Why did the second wave of globalization lead to political fragmentation? To answer these questions, this paper develops a theoretical framework to study the interaction between globalization and political structure. We view globalization as a process by which the geographical size of markets expands. There is wide consensus that globalization started around the mid nineteenth century and that it was fueled by major technological advancements, sustained

¹See Appendix A.1 for more details about the data.

²Most importantly, the commonly used “International System” (Singer and Small 1966), including countries with international recognition, grossly underestimates the number of independent political entities in the 19th-century developing world.

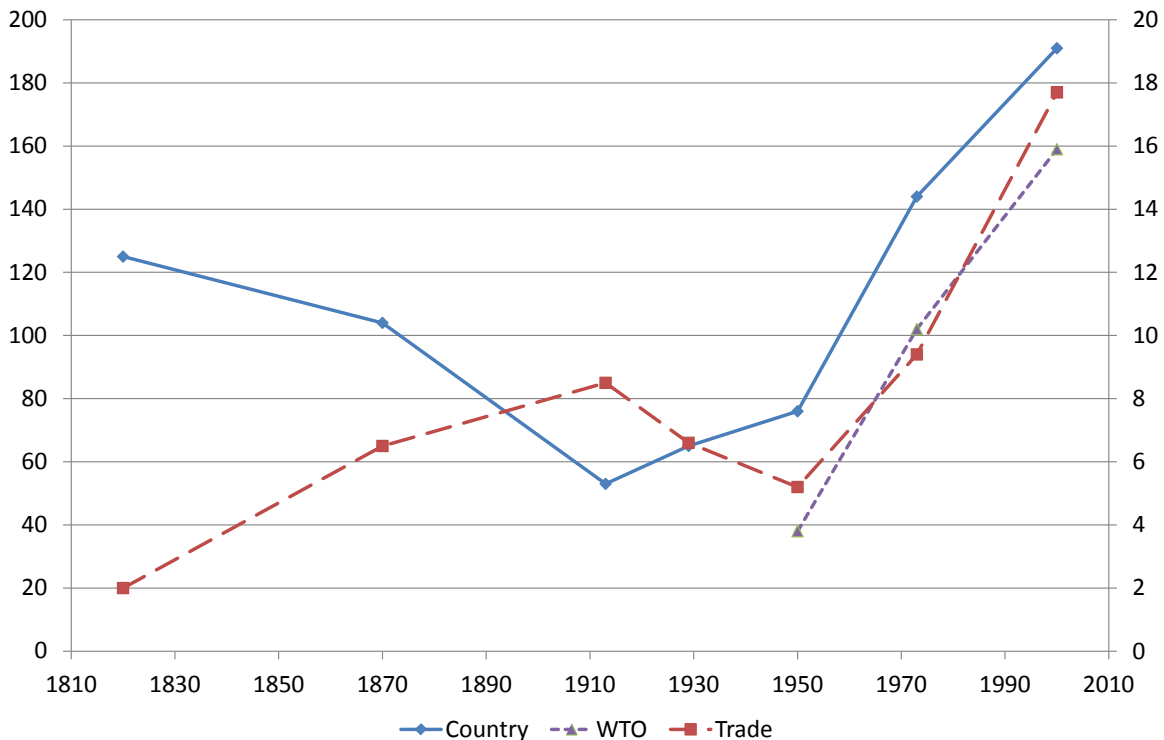


Figure 1: Trade share (right axis), the number of countries and WTO membership (left axis.) See appendix for details on data.

economic growth and the adoption of market-promoting policies.³ A key premise behind our work is that borders hamper trade, and globalization makes borders more costly. Consistent with the evidence in Figures 1 and 2, our theory suggests that political structure adapts to expanding trade opportunities in a non-monotonic way. In its early stages, the political response to globalization consists mostly of removing borders by increasing the size of countries. In its later stages, however, the political response to globalization is to remove borders by creating economic unions, and this leads to a reduction in the size of countries.

Our theory builds on the classic work of Alesina and Spolaore (1997), who argued that country size trades off preference heterogeneity and economies of scale. Preference heterogeneity across space creates a force towards political fragmentation. Each relevant geographical unit or “locality” would like to have a government type that is as close to its ideal type as possible. As the size of a country increases by adding more localities, the resulting government type tends to move farther away from the ideal type of each of them. Thus, the mismatch between citizens’ ideal and actual government type increases with country size.

³See Hugot (2015) and Pascali (2014) for recent papers on the onset of globalization.

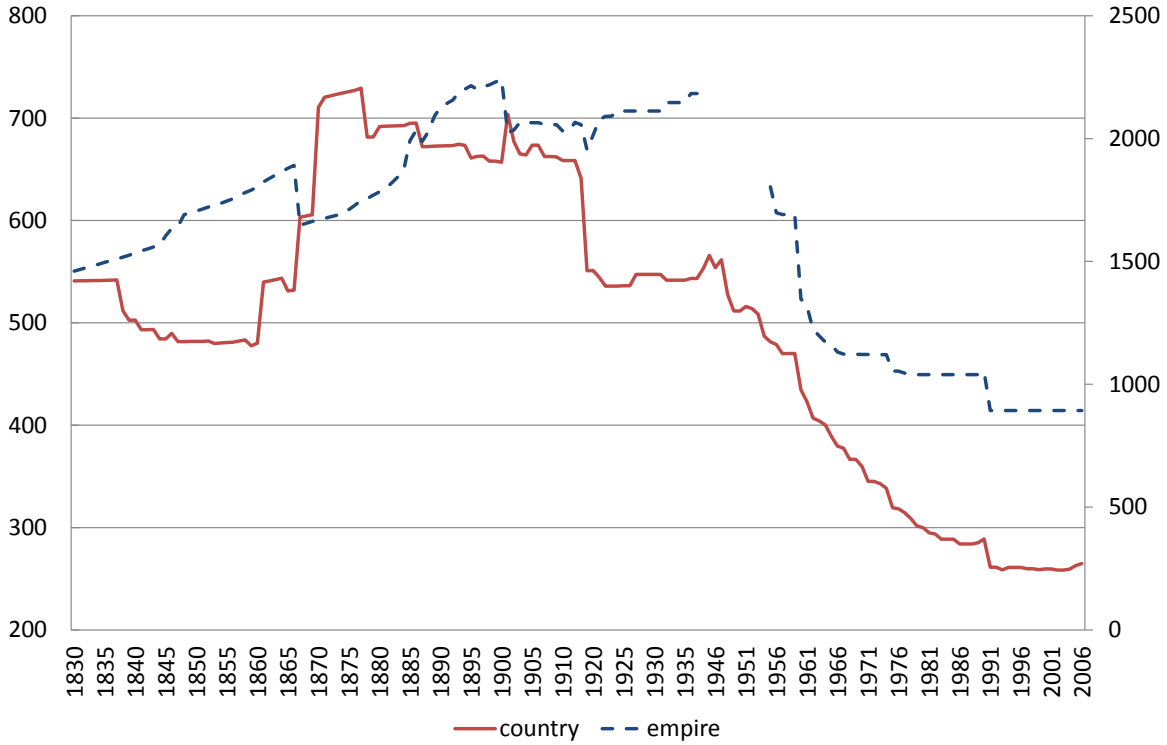


Figure 2: Size of countries (left axis) and empires (right axis). Thousand squared miles, see the appendix for details on data.

Economies of scale create a force towards political concentration. As country size increases, some fixed costs of setting up and running a government are shared by a growing number of localities. Thus, the cost of the government also declines with size. Alesina, Spolaore and Wacziarg (2000) extended this theory to allow for international trade and a border effect. This creates another force for political concentration. As country size increases, the number of borders declines and the gains from trade grow. Then, they modeled globalization as a reduction in the border effect, which weakens the benefits of size and leads to political fragmentation.⁴

Our theory departs from this earlier work in two important ways. The first one is our approach to globalization. Instead of modeling globalization as a reduction in the border effect, we model this phenomenon as an increase in trade opportunities. This is more consistent with the notion of globalization as a geographical extension of the market. Many goods and

⁴In a similar vein, but less directly related to our work, Casella (2001) and Casella and Feinstein (2002) study the interaction between economic and political integration in a model where public goods affects the gains from trade, while Bolton and Roland (1996, 1997) focus on the role of income distribution and factor mobility. See Ruta (2005) for a survey of the literature on the economic incentives for countries to integrate or separate, and how these incentives are affected by globalization.

services that were traded only locally in the early nineteenth century were already traded internationally by the early twentieth century. There are even more that are traded internationally in the early twenty-first century. Globalization effectively makes borders more costly, as they affect a growing volume of trade. This change of perspective suggests that globalization provides incentives to remove borders rather than to create them, and it would lead us to the conclusion that globalization leads to political concentration.

The second and most important novelty of our theory is to recognize that the notion of government type is multidimensional. In particular, this notion includes a full description of the laws and institutions that govern economic and cultural interactions within a country. By economic interactions, we mean market exchanges of any sort. By cultural interactions, we mean “the distinctive ideas, customs, social behavior, products, or way of life of a particular nation, society, people, or period.”⁵ Thus, it is useful to think about the economic and cultural types of government.⁶

This distinction enriches the analysis since it expands the set of possible political structures. We can think of countries as sets of localities that share the same economic and cultural types of government. We can then think of economic unions as sets of countries that share the same economic type of government but not the same cultural type. Going back to Figures 1 and 2, we can now ask: Why did the first wave of globalization reduce the number of countries but did not generate economic unions? Why did the second wave of globalization increase the number of countries and led to the creation of economic unions? Once these questions have been re-formulated this way, we need only two additional (and reasonable) assumptions to obtain our main result.

The first assumption is that the border effect is largely due to differences in the economic type of government, while preference heterogeneity is defined mostly over the cultural type of government. This suggests the benefits of a political structure consisting of small countries each of them with its own distinct cultural type, and a large economic union that encompasses many or all of these countries. This political structure would allow localities to enjoy a

⁵This is the definition of culture offered by the *Oxford English Dictionary*. This notion of a group’s culture is quite broad. It includes not only sharing a common language and history, but also sharing common views about justice and punishment, how to educate children, the size and scope of the welfare system, the need for redistribution, and so on.

⁶Thus, our research is also related to the broad theory of fiscal and political federalism. The classic economic analysis of federalism (e.g., Oates 1972) focuses on two rationales for centralization: economies of scale and benefits from policy coordination. Political-economy analyses have emphasized that a countervailing cost of centralization is the reduced ability to tailor policy choices to local tastes (Lockwood 2002; Besley and Coate 2003; Boffa, Piolatto and Ponzetto 2015). We study the evolution of this trade-off in the wake of globalization and its implication for the endogenous formation of countries and economic unions.

government that is close to their ideal type. It would also facilitate trade by removing borders.

The second assumption is that there are limited economies of scope when the cultural and economic types of government coincide. Having different economic and cultural types is akin to having two different sets of laws and institutions for market and cultural interactions or, in short, having two different levels of government. This not only duplicates costs, but it also introduces transaction costs across levels of government. Earlier research, and most notably the work of Alesina et al., has adopted the view that economies of scope are so large that the possibility of having different economic and cultural types of government is not worth considering. We relax this assumption and allow for limited economies of scope.

The non-monotonic effect of globalization on political structure follows from these assumptions. At early stages of globalization, localities take advantage of increased trade opportunities by enlarging country size. That is, the cost of reaping the gains from trade is paid in terms of a larger mismatch between the ideal and actual types of government. Forming economic unions is not worthwhile since trade is small and the benefits of removing borders this way are outweighed by loss of economies of scope. As globalization proceeds, the size of countries increases and the cost of preference mismatch keeps growing. Eventually, this cost is too large and it is worth giving up economies of scope. The world political structure shifts from a world with a few large countries to a new world with many small countries that participate in an economic union. This result is consistent with the observation that political structure reacted differently to the first and second waves of globalization.

A key aspect of any theory of political structure is a view of how localities interact. In Section 2, we develop a stylized model of the world with a continuum of localities that are sovereign and free to sign political contracts among themselves. Such a view of political interactions can aptly be described as the rule of diplomacy. We show that equilibrium political contracts take the form of countries and unions. And we describe how the number and size of countries and unions evolves with globalization. In particular, we find that in the early stages of globalization, country size grows and there are no economic unions; while at the later stages of globalization, a world economic union is formed and country size shrinks.

But diplomacy is hardly the only form of interaction in the real world. In Section 3, we allow a small set of great powers to wage war and build empires. In equilibrium, the world is partitioned into a set of empires ruled by war and conquest and a free world ruled by diplomacy. Interestingly, we find that our main result also applies to this world: in the early stages of globalization, country size grows and there are no economic unions; while at later

stages of globalization, a world economic union is formed and country size shrinks. In both stages, great powers choose not to build empires and instead join the free world. Diplomacy rules. The new result now is that there exists an intermediate stage of globalization in which great powers build empires that grow in size. These empires eventually collapse as the world enters the third and final stage of globalization and the world economic union makes them obsolete.⁷

In Section 4, we break the perfect symmetry of the benchmark model by assuming that the world is divided into regions and that trade costs and cultural differences vary across and within regions. We then use the model to rationalize why the process of formation of unions tends to be gradual and often starts at the regional level. The picture that arises from our stylized model is therefore rich and suggestive. In Section 5 we show how our analytical results help us interpret a variety of specific historical experiences. Section 6 concludes.

2 THE SYMMETRIC WORLD

In this section, we develop a stylized model of the world that contains the basic ingredients of the theory: geography, markets and preferences. The model mixes these ingredients imposing a high degree of symmetry. This allows us to derive our basic results on the effects of globalization on political structure quickly and intuitively.

The concept of locality is a key primitive in our theory. We model the world as a set of locations within which there are neither geographical nor cultural distances, and we label them localities. Thus, localities consist of a group of people sharing common preferences and inhabiting a particular territory. This approach, which is common in the literature, simplifies the study of how peoples with different cultures interact and organize themselves into political entities. But it is silent about how these cultural differences arose in the first place and how they evolve over time. It also abstracts from domestic conflict.

The concept of globalization is another important primitive in our theory. Geographical distances introduce trade costs across localities that shelter local markets. We interpret globalization as exogenous technological change that gradually removes these trade costs and turns local markets into global ones. In particular, we use the popular traded-nontraded model in which the markets for traded goods are global, while the markets for nontraded goods are local. We then model globalization as an increase in the share of traded goods.

⁷See Martin, Mayer and Thoenig (2008) for a theoretical and empirical analysis of the relationship between trade and war.

2.1 BASIC SETUP

We consider a world with a continuum of atomistic geographical units called localities, $l \in [0, 1]$. Each locality contains a positive measure of identical individuals. Let W_l be the welfare of the representative individual of locality l . For short, we refer to this individual as “locality l .” Then, the welfare of locality l is determined as follows:

$$W_l = W_l^M + W_l^G, \quad (1)$$

where W_l^M is the utility derived from the consumption of market goods, and W_l^G is the utility derived from the services provided by its government. We explain next how these two variables are determined.

2.1.1 Markets and Trade

There is a continuum of industries producing goods, $i \in [0, 1]$. Let $c_l(i)$ be the consumption of goods of industry i by locality l . The utility function takes the following form:

$$W_l^M = \int_0^1 \ln c_l(i) di. \quad (2)$$

The choice of a symmetric logarithmic or Cobb-Douglas utility function ensures that all localities spend the same share of their income in each industry.

The production of consumption goods requires differentiated input varieties, $m \in [0, 1]$. Define $c_l(m, i)$ as the amount of inputs of the variety m used by locality l in the production of goods in industry i . Then, we have that:

$$c_l(i) = \exp \left\{ \int_0^1 \ln c_l(m, i) dm \right\}. \quad (3)$$

This production function is symmetric across and within industries. Since we use again the convenient Cobb-Douglas formulation, each locality spends the same fraction of its income on all varieties of all industries.

To introduce gains from specialization and trade, we use a simple symmetric version of the Ricardian model. Each locality is endowed with one unit of labor in each industry. This unit can produce one unit of the variety with the same index as the locality, i.e. $m = l$; or $e^{-\eta}$ units of any other variety, i.e. $m \neq l$. Since $\eta > 0$, each locality has a technological advantage in its “own” variety. The parameter η measures the extent to which technologies

differ across localities and, therefore, the potential gains from specialization and trade.

We introduce both technological and policy-induced barriers to trade. In some industries, the costs of shipping inputs across localities are negligible, while in the rest these costs are prohibitive. As usual, we refer to these two sets of industries as traded and nontraded, respectively. We order industries so that traded ones have low indices, i.e., $i \in [0, \tau]$. We interpret globalization as an increase in τ , which reflects the decline of technological trade costs.

The presence of borders also imposes costs to trade. Define a variable b_{lm} which takes value β if there is a border between localities l and m , and it takes value 0 otherwise. Then, we assume that only a fraction $e^{-b_{lm}}$ of the goods shipped from l to m arrive to their destination. Thus, the parameter β measures the size of the border effect. To ensure that borders do not preclude specialization and trade, we assume that $\eta > \beta > 0$.

A market equilibrium is a set of prices and quantities such that individuals maximize utility and markets clear. Appendix A.2 shows that there exists a unique market equilibrium for each possible configuration of borders. In this equilibrium, traded industries specialize in the locality's input variety, export essentially all of their production and import the remaining input varieties. Thus, consumption in traded industries is $c_l(m, i) = e^{-b_{lm}}$ if $i \in [0, \tau]$. Nontraded industries are forced to produce locally all input varieties. Thus, consumption in nontraded industries is $c_l(m, i) = e^{-\eta}$ if $i \in (\tau, 1]$. This implies the following utility from consuming market goods:

$$W_l^M = -(1 - \tau)\eta - \tau \int_0^1 b_{lm} dm. \quad (4)$$

Equation (4) is valid for any configuration of borders. An increase in the share of traded goods raises the consumption of newly traded goods. The size of this gain is proportional to η for trade that does not cross a border, and it is proportional to $\eta - \beta$ for trade that crosses a border. Removing borders also raises consumption. The size of this gain equals $\tau\beta$ per border removed. These results are intuitive and play an important role in the analysis.

2.1.2 Government Type

There is a continuum of basic government varieties, $x \in [0, 1]$. A government for locality l consists of a pair of density functions $g_l^E(x)$ and $g_l^C(x)$ defined over these varieties. Since these functions are densities, $g_l^E(x) \geq 0$, $g_l^C(x) \geq 0$ and $\int_0^1 g_l^E(x) dx = \int_0^1 g_l^C(x) dx = 1$. We refer to the functions g_l^E and g_l^C as the economic and cultural types of the government

of locality l , respectively. These types constitute a complete description of the laws and institutions that govern economic and cultural interactions within locality l .

We now introduce three assumptions about government services. The first is that each locality has a different ideal cultural variety of government. We define and order the basic varieties such that the ideal one for locality l is $x = l$. The value of government services for locality l depends positively on how dense g_l^C is at $x = l$, that is, it depends on $g_l^C(l)$.

The second assumption is that there are economies of scale in the provision of government services. Building and maintaining a specific government reduces the overall provision of government services by a total amount $\phi > 0$. Localities that choose the same pair (g_l^E, g_l^C) share this cost equally among them.

The third assumption is that there are economies of scope when cultural and economic types of government coincide. Localities that choose $g_l^C \neq g_l^E$ must pay a per capita cost of $\kappa > 0$.

To formalize these assumptions, we write the utility derived from government services as follows:

$$W_l^G = -\frac{\delta}{g_l^C(l)} - \frac{\phi}{\int_0^1 I_{g_l=g_m} dm} - \kappa I_{g_l^C \neq g_l^E}, \quad (5)$$

where $\delta > 0$; $I_{g_l=g_m}$ is an indicator variable that takes value one if $(g_l^E, g_l^C) = (g_m^E, g_m^C)$, and zero otherwise; and $I_{g_l^C \neq g_l^E}$ is another indicator variable that takes value 1 if $g_l^E \neq g_l^C$, and zero otherwise. The first term in Equation (5) says that the value of government services for locality l grows with the density of its own specific ideal variety in the cultural type of its government. The parameter δ measures the importance of preference heterogeneity. The second term in Equation (5) says that the fixed cost of government declines with the number of localities that share the same government. The parameter ϕ measures the magnitude of these economies of scale. The third term of Equation (5) says that choosing different economic and cultural types of government reduces the value of government services. The parameter κ measures the magnitude of these economies of scope.

We next introduce a fourth assumption, namely, that having different economic types creates borders that obstruct trade:

$$b_{lm} = \beta I_{g_l^E \neq g_m^E}, \quad (6)$$

where $I_{g_l^E \neq g_m^E}$ is an indicator variable that takes value 1 if $g_l^E \neq g_m^E$, and zero otherwise. The idea, of course, is that disparity in economic laws and institutions makes trade more costly. The parameter β is a measure of these costs.

2.2 EQUILIBRIUM POLITICAL STRUCTURE UNDER DIPLOMACY

A political structure for the world is a set of governments for all localities, i.e. $\{g_l^C, g_l^E\}_{l \in [0,1]}$. Combining Equations (1), (4), (5) and (6), we obtain:

$$W_l = -(1 - \tau)\eta - \tau\beta \int_0^1 I_{g_l^E \neq g_m^E} dm - \frac{\delta}{g_l^C(l)} - \frac{\phi}{\int_0^1 I_{g_l = g_m} dm} - \kappa I_{g_l^C \neq g_l^E}. \quad (7)$$

Equation (7) shows how political structure determines welfare and reveals the key trade-off that underlies our theory. A desirable political structure should facilitate trade, accommodate preference heterogeneity and take advantage of economies of scale and scope. But these goals cannot be achieved simultaneously and something must give.

To determine the equilibrium political structure, we must make assumptions about how governments are chosen. A good starting point is the case in which localities are sovereign and they freely choose their own economic and cultural types of government. Since there is no heterogeneity within localities, their choices maximize the welfare of their representative individual. Their choice of cultural type does not affect other localities. But their choice of economic type does, since it determines borders and therefore trade costs. These policy externalities provide incentives for localities to engage in diplomacy and negotiate political contracts.

Localities can sign political contracts to share a government with other localities. We refer to these contracts and the localities that sign them as countries. Formally, a country is a set of localities that share the same economic and cultural types of government, i.e. $(g_l^C, g_l^E) = (g_m^C, g_m^E)$ for all l and m belonging to the country. Countries choose the average ideal cultural type of their members:

$$g_l^C(x) = \begin{cases} \frac{1}{S} & \text{if } x \in \text{country} \\ 0 & \text{if } x \notin \text{country} \end{cases}, \quad (8)$$

where S is the combined measure of localities that belong to the country.

Countries can also sign political contracts to share a common economic type of government. We refer to these contracts and the countries that sign them as ('economic') unions. Formally, a union is a set of countries that have different cultural types of government, but share the same economic type of government, i.e. $g_l^E = g_m^E$ for all l and m belonging to the

union.⁸ Unions choose the average ideal type of their members:

$$g_l^E(x) = \begin{cases} \frac{1}{U} & \text{if } x \in \text{union} \\ 0 & \text{if } x \notin \text{union} \end{cases}, \quad (9)$$

where U is the combined measure of localities that belong to the union. As a result, $g_l^C \neq g_l^E$ for all l belonging to the union.⁹

How does the world organize itself into countries and unions? A natural benchmark is that in which bargaining among localities is efficient and delivers Pareto efficient outcomes. In this case, the equilibrium political structure under diplomacy is obtained by solving this maximization problem:

$$\{g_l^C, g_l^E\}_{l \in [0,1]} = \arg \max \int_0^1 \omega_l W_l dl, \quad (10)$$

where $\{\omega_l\}_{l \in [0,1]}$ is a set of Pareto weights such that $\int_0^1 \omega_l dl = 1$. Given all the symmetry of this world, it seems reasonable to focus on the case in which the bargaining process treats all localities in the same way, i.e. $\omega_l = 1$. Sometimes this political structure is referred to as the utilitarian welfare optimum.

To find this political structure, we proceed in three steps. First, we compute the welfare W_N generated by a world without unions:

$$W_N = \max_{S_N \in [0,1]} \left\{ -(1-\tau)\eta - \tau\beta(1-S_N) - \delta S_N - \frac{\phi}{S_N} \right\}, \quad (11)$$

where S_N is the preferred country size in the absence of unions.¹⁰ This political structure takes full advantage of economies of scope, and country size trades off preference heterogene-

⁸We could also define ‘cultural’ unions analogously. But this is unnecessary since countries would never sign political contracts to share a common type of cultural government, and then choose different economic types of government.

⁹We assume that both countries and unions choose the average ideal type of their members for simplicity. But it is not difficult to derive this as a result. No additional assumptions are needed to show that countries choose the average cultural and economic type of their member localities. To show that unions choose the average economic type of their member countries, we would need to assume some preference heterogeneity also for economic type.

¹⁰Throughout, we disregard the constraint that the number of countries, which in this case is $1/S_N$, must be a natural number. Aside from this constraint, all localities prefer the same optimal country size, so any Pareto-efficient equilibrium is symmetric. In Appendix A.3 we introduce this integer constraint and show that the utilitarian welfare optimum remains symmetric and qualitatively analogous to the tractable approximation we present here.

ity against both economies of scale and facilitating trade:

$$S_N = \sqrt{\frac{\phi}{\delta - \tau\beta}}. \quad (12)$$

The equilibrium size of countries in the absence of unions is increasing with economies of scale (ϕ) and the importance of trade ($\tau\beta$), and it is decreasing with preference heterogeneity (δ). Equation (12) assumes that $\delta > \phi + \beta$, so that there is enough preference heterogeneity to ensure that countries are always smaller than the whole world.

Second, we compute the welfare W_U generated by a world with unions. The cost of forming a union does not change with its size, but its benefit does increase with size. Thus, the equilibrium union is always the world union, i.e. $U = 1$. This allows us to write:

$$W_U = \max_{S_U \in [0,1]} \left\{ -(1 - \tau)\eta - \delta S_U - \frac{\phi}{S_U} - \kappa \right\}, \quad (13)$$

where S_U is the preferred country size with a world union. This political structure gives up economies of scope in order to remove borders and facilitate trade. Then, country size trades off preference heterogeneity and economies of scale:

$$S_U = \sqrt{\frac{\phi}{\delta}}. \quad (14)$$

The equilibrium size of countries with a world union is increasing with economies of scale (ϕ) and it is decreasing with preference heterogeneity (δ). Country size is always smaller with a world union than without. The reason is that the union removes one of the incentives for country size, namely, facilitating trade.

The third step is to determine the equilibrium political structure as follows:

$$W_F = \max \{W_N, W_U\}, \quad (15)$$

where W_F is the welfare obtained by a locality in this free world. If $W_N > W_U$, the equilibrium political structure consists of the world being partitioned into countries of size S_N . If instead $W_N < W_U$, the equilibrium political structure consists of the world being partitioned into countries of size S_U and a world union. Naturally, in the knife-edge case in which $W_N = W_U$, both solutions are equilibrium political structures.

A little bit of algebra shows that:

$$W_F = -(1 - \tau)\eta - \tau\beta - \min \left\{ 2\sqrt{\phi(\delta - \tau\beta)}, 2\sqrt{\phi\delta} + \kappa - \tau\beta \right\}. \quad (16)$$

It is straightforward to see that the world union is preferred if and only if $2\sqrt{\phi\delta} + \kappa - \tau\beta < 2\sqrt{\phi(\delta - \tau\beta)}$, hence for high values of τ , β and δ ; and low values of ϕ and κ . A world union is more useful if trade costs are low, the border effect is sizeable and there is substantial preference heterogeneity. A world union is less useful if economies of scale and scope are sizable.

2.3 GLOBALIZATION AND POLITICAL STRUCTURE

With these results at hand, we can now return to Figures 1 and 2 in the introduction and ask again: Why did the first wave of globalization reduce the number of countries but did not generate economic unions? Why did the second wave of globalization increase the number of countries and led to the creation of economic unions?

To develop some intuition, Figure 3 plots W_N , W_U and W_F as a function of τ . Note that W_F is the upper envelope of W_N and W_U , and it is depicted as a solid line. As τ increases, the potential gains from trade increase and so does the incentive to remove borders. This affects political structure in a non-monotonic way. In particular, there is a threshold level τ_U such that $W_N < W_U$ and there is no world union for $\tau < \tau_U$; while $W_N > W_U$ and there is a world union for $\tau > \tau_U$. This threshold is interior, i.e. $\tau_U \in (0, 1)$, for intermediate levels of economies of scope:

$$0 < \kappa < \beta - 2\sqrt{\phi} \left(\sqrt{\delta} - \sqrt{\delta - \beta} \right). \quad (17)$$

If economies of scope are negligible, the smallest amount of trade leads to the formation of a world union. If economies of scope are prohibitive, the world union is never an equilibrium. The comparative statics of this threshold follow directly from our analysis of Equation (16). The larger is the border effect (β) and preference heterogeneity (δ), the smaller is τ_U . The larger are economies of scale (ϕ) and scope (κ), the larger is τ_U .

Figure 4 shows how the equilibrium political structure changes with globalization by plotting the equilibrium size of countries and unions. At low levels of globalization ($\tau \leq \tau_U$), expanding trade opportunities lead to an increase in country size. At this early stage of globalization, the cost of reaping new gains from trade is a growing preference mismatch.

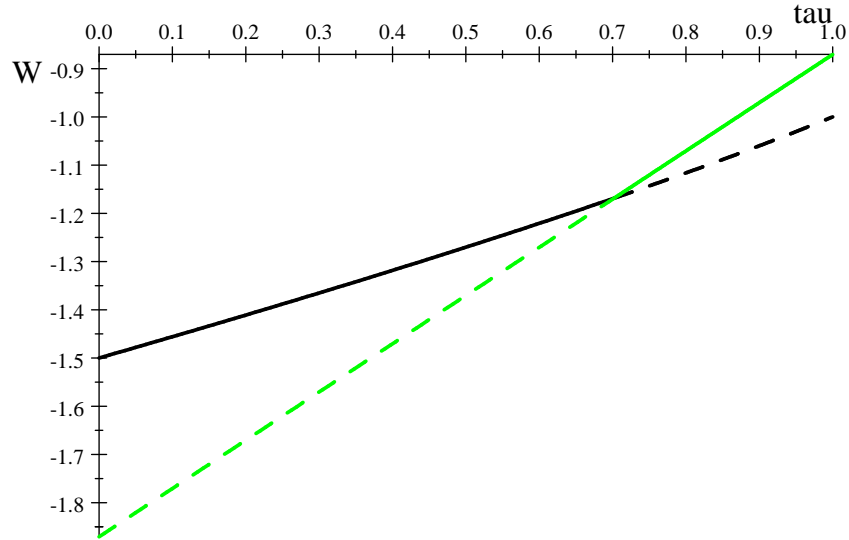


Figure 3: Globalization and Welfare. The figure shows how equilibrium welfare (W_F) changes with globalization (τ). The black line is maximum welfare without a union (W_N), the green line with the world union (W_U).

It is simply too expensive to create a world union given the small amount of trade. At high levels of globalization ($\tau \geq \tau_U$), the preference mismatch has grown too large and expanding trade opportunities now lead to the creation of a world union. At this late stage of globalization, the cost of reaping new gains from trade is the loss of economies of scope. Indeed, the creation of a world union allows countries to revert to a smaller size and reduce the preference mismatch. Further increases in τ have no effect on political structure.

The symmetric world neatly illustrates the non-monotonic response of political structure to globalization. It is instructive to compare this result to the finding by Alesina, Spolaore and Wacziarg (2000) that globalization always reduces country size. Their analysis differs from ours in two crucial ways. The first one is that, unlike us, Alesina et al. assumed that κ is prohibitive so that the world union is never an equilibrium and country size is determined by Equation (12). The second difference is that, unlike us, Alesina et al. interpreted globalization as a reduction in the border effect β , keeping constant the fraction of traded goods τ (which was set to one). It follows directly from Equation (12) that a reduction in the border effect reduces country size. In our model, the border effect is removed endogenously by the creation of the union, but only when τ is sufficiently high.

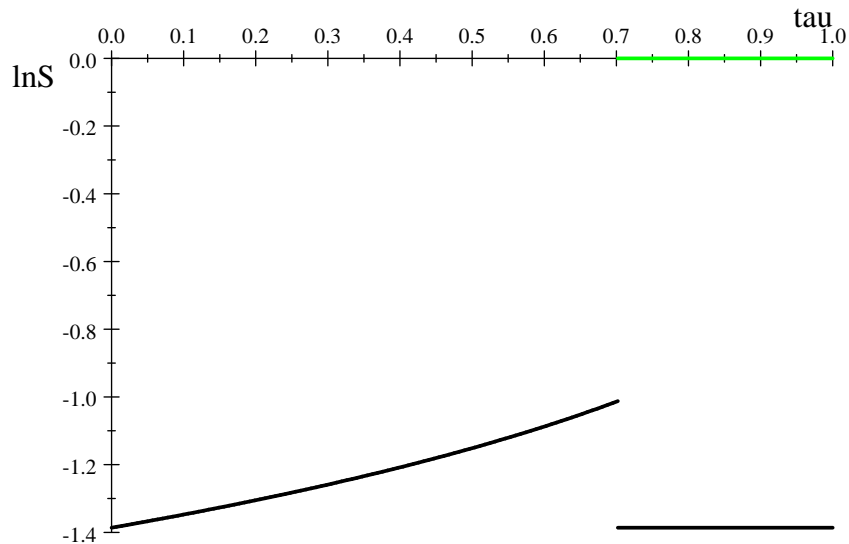


Figure 4: Globalization, Countries and Unions. The figure shows how the world political structure changes with globalization (τ). The black line is the size of each country, the green line is the world union.

3 WAR AND CONQUEST

In our baseline model we considered a world in which every locality can choose its own government freely. Then, the equilibrium political structure is the outcome of free contracting among equals. This constitutes a useful theoretical benchmark and an ideal state of affairs. But it hardly reflects historical experience. Alongside diplomacy, war and conquest (or the threat of it) have played a crucial role in shaping real-world political structure. Thus, we want to know whether the possibility of war and conquest affects the relationship between globalization and political structure. And if it does, we also want to know how and why this happens.

We keep all assumptions as before regarding preferences and technology. But now we consider an alternative rule to determine political structure. In particular, we assume that the world contains “great powers” that are able to conquer other localities through war and form empires. If some of these great powers choose to do so, the world is partitioned into a set of empires where political structure is determined by war and conquest, and a set of free countries where political structure is determined by diplomacy. We explain how this all works next.

3.1 EMPIRE-BUILDING

The world contains a finite set of $\pi \in \mathbb{N}$ identical “great powers” that are able to conquer other localities through war and form empires. Like a free country, an empire consists of a set of localities that share the same economic and cultural types of government, i.e. $(g_l^C, g_l^E) = (g_m^C, g_m^E)$ for all l and m belonging to the empire. Unlike a free country, an empire does not treat all of its member localities symmetrically. In particular, an empire contains two distinct sets of localities: the metropolis and the colonies. Let E be the size of the empire containing a metropolis of size M and colonies of size $E - M$. The key asymmetry is that the empire’s cultural type is the average ideal cultural type of the metropolis:

$$g_l^C(x) = \begin{cases} \frac{1}{M} & \text{if } x \in \text{metropolis} \\ 0 & \text{if } x \notin \text{metropolis} \end{cases} . \quad (18)$$

Since this implies that colonies have a welfare of $W_C = -\infty$, no locality becomes a colony voluntarily.

To build an empire of size E , a great power must find allies to build a metropolis, and then wage war to conquer the colonies. A colonial war is successful if and only if the size of the metropolis is large enough relative to that of the colonies:

$$M \geq \mu E. \quad (19)$$

The parameter $\mu \in (0, 1)$ provides the smallest size of a successful metropolis.

From the perspective of the metropolis, the upside of building an empire is that it forces the colonies to adopt the average ideal government type of the metropolis. This facilitates trade and generates economies of scale without creating preference mismatch. The downside of building an empire is that waging war and holding the empire together reduces the utility that the metropolis derives from government services by an amount $\omega > 0$. This cost captures the diversion of government resources from providing public goods in the metropolis to waging colonial wars.

Localities that do not belong to empires remain sovereign and choose their political structure through diplomacy. In particular, they form free countries as in the baseline model.

3.2 EQUILIBRIUM POLITICAL STRUCTURE WITH WAR AND CONQUEST

We now determine equilibrium political structure arising from the following three stages:

1. The great powers choose whether to wage war and build empires. Other localities might become their allies, their colonies, or remain sovereign.
2. Localities that remain sovereign form countries by free and symmetric contracting, as in our baseline model.
3. Empires and free countries can form economic unions.

We assume that great powers can successfully coordinate with each other on the creation of their respective empires. Coordination ensures that great powers wage colonial wars without ever clashing with each other. They agree on their separate spheres of influence, so they can target disjoint sets of localities to conquer as their colonies. Moreover, the number of great powers is small enough that all of them can find the desired number of localities to be used as allies and colonies without having to fight each other.¹¹

To solve for the equilibrium, we first study the incentives of great powers to build empires. Given the choice of great powers, we then study the equilibrium political structure of the remaining sovereign localities.

3.2.1 *The great powers*

Great powers must first decide whether to build an empire and become part of the metropolis or to forego war. To do this, they compare the welfare of these two options. If the great powers agree to refrain from waging war and instead jointly choose to form countries and unions by free and symmetric contracting, the world equilibrium remains the one described by the previous section. All localities, including the great powers, have welfare W_F , which is still given by Equation (16).

However, great powers may decide to wage war and form empires. In this case, the welfare W_E of a metropolis in an empire outside of an economic union is:

$$W_E = \max_{E \in [0,1]} \left\{ -(1 - \tau)\eta - \tau\beta(1 - E) - \delta M - \frac{\phi}{E} - \omega \right\} \text{ s.t. } M \geq \mu E. \quad (20)$$

¹¹Historically, great powers have in fact agreed on such explicit partitions of the world, from the Treaty of Tordesillas in 1494 to the Berlin Conference in 1884.

Naturally, the constraint is binding and $M = \mu E$. From the perspective of the metropolis, there is no reason to convert additional colonies into allies and “worsen” the government type. The equilibrium size of the empire trades off preference heterogeneity against both economies of scale and facilitating trade:

$$E = \sqrt{\frac{\phi}{\delta\mu - \tau\beta}}. \quad (21)$$

The equilibrium size of empires is increasing with economies of scale (ϕ) and the importance of trade ($\tau\beta$), and it is decreasing with preference heterogeneity (δ). These comparative statics are the same as for countries. Also, the smaller is the size of the metropolis relative to that of the colonies (μ) the larger is the empire. We assume that $\delta\mu > \phi + \beta$ such that an empire never grows as large as the entire world.

Comparing Equation (21) to Equation (12), it is immediate to see that empires are larger than peaceful countries, because the metropolis does not internalize the cost of the cultural mismatch imposed on the colonies. Moreover, comparing W_E to W_F (i.e., Equations 20 and 16) in the case $\tau \rightarrow 0$ shows that diplomacy prevails in autarky if war is sufficiently costly:

$$\omega > 2\sqrt{\phi\delta}(1 - \sqrt{\mu}). \quad (22)$$

This condition implies that great powers do not wish to conquer colonies merely to compel them to defray the fixed cost of government. The motivation for imperial expansion lies instead in the desire to gain access to colonial markets without having to compromise the great power’s preferred type of cultural government. As a consequence, the great powers choose to forego warfare in autarky, when market access is worthless. When this condition is satisfied, war may arise endogenously as a function of globalization, but it is not the dominant strategy for any value of τ . To focus on this more interesting and realistic case, from now on, we assume that Condition (22) is satisfied.

Consider now an empire part of an economic union. Since the benefit of forming a union is increasing in its size, great powers will jointly agree to form the largest possible union. Moreover, it is straightforward to show that when great powers want to form a world union the remaining localities are always willing to join it.¹² Hence, if the union is chosen, it will be the world union. However, an additional implication of Condition (22) is that the great

¹²Intuitively, since free countries are always smaller than empires, they are always keener on joining a world union and reaping the benefits of reducing barriers to international trade.

powers prefer to forego war if they wish to join an economic union that guarantees unfettered access to world markets. Hence, economic unions including great powers will be peaceful.

In sum, welfare for great powers is $W_P = \max\{W_E, W_F\}$. When $W_E > W_F$, there are π empires of size E .¹³ When $W_E \leq W_F$, instead, diplomacy rules. While we know already that peace prevails in autarky, great powers may choose to wage war at intermediate levels of globalization. To see this, note that:

$$\frac{\partial}{\partial \tau} (W_E - W_N) = \beta (E - S_N) > 0. \quad (23)$$

Rising globalization makes empires more appealing relative free countries because in equilibrium they have larger markets. For the same reason, however, globalization also makes a peaceful world union more appealing relative to an empire.

3.2.2 *The free world*

Consider now localities that are neither great powers nor part of an empire. If $W_F \geq W_E$, as we already discussed above, all localities have welfare W_F and the world political structure is identical to the case without war. If $W_F < W_E$, instead, there are π empires of size E . The analysis of the remaining $1 - \pi E$ sovereign localities, i.e., the “free world”, is quite similar to that of the previous section. The welfare delivered by a free world without unions is still described by Equation (11), and the associated equilibrium country size S_N is still given by Equation (12). There is however a new possibility. All localities that do not belong to empires may form a union. If they choose to do so, it is again the largest possible one. But now this is the union of all free countries in a world with empires. Since this union has a combined size of $1 - \pi E$, the welfare of its localities is given by:

$$\check{W}_U = \max_{S_U \in [0,1]} \left\{ - (1 - \tau) \eta - \tau \beta \pi E - \delta S_U - \frac{\phi}{S_U} - \kappa \right\}. \quad (24)$$

Empires reduce the size of the free world and therefore the welfare delivered by an economic union of free countries. Note however that equilibrium country size S_U is still described by Equation (14).

If $W_N > \check{W}_U$, the free world is partitioned into countries of size S_N and there is no economic union. If instead $W_N < \check{W}_U$, the free world is partitioned into countries of size S_U

¹³We assume that great powers coordinate on forming the maximum number of empires rather than joining each other’s metropolises. This is feasible given our assumption $\pi E < 1$ for all τ .

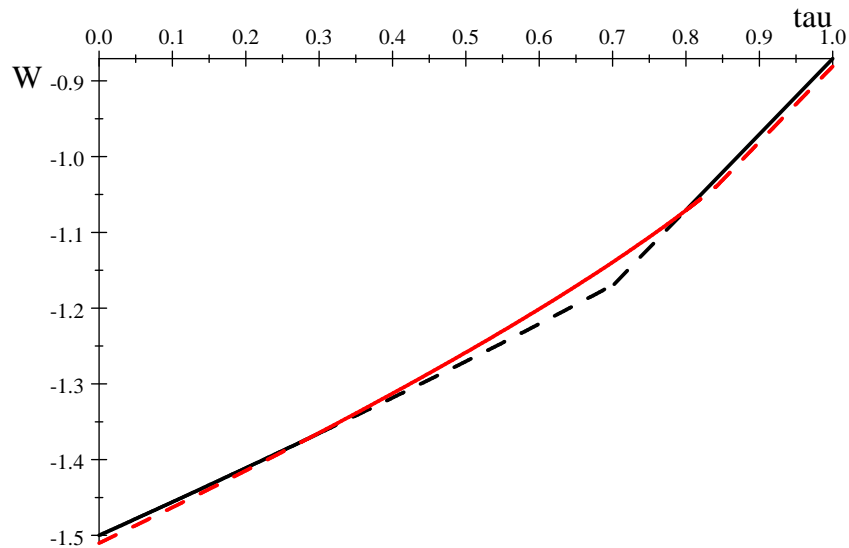


Figure 5: Contracts, Conquest and Welfare. The figure shows how equilibrium welfare changes with globalization (τ). The black line is maximum welfare from contracts (W_F), the red line from conquest (W_E).

and there is a union. Comparing Equation (24) to Equation (11), it is easy to see that the union of free countries is preferred if trade costs are low (τ), the border effect is sizeable (β), there is substantial preference heterogeneity (δ), and economies of scale (ϕ) and scope (κ) are weak. These results were already obtained in the absence of empires, i.e. with $\pi E = 0$. A new result here is that empires (πE) reduce the value of an economic union. The reason, of course, is that they reduce its size.

3.3 GLOBALIZATION AND POLITICAL STRUCTURE AGAIN

Let us now return to the analysis of the effects of globalization on political structure that motivates our analysis. Figure 5 plots W_F and W_M as a function of τ . In the most interesting case that we plot, there are now two relevant threshold values, τ_E and τ_U . These thresholds define three different stages of globalization with distinct equilibrium political structures.

Figure 6 depicts this three-stage evolution by plotting the equilibrium size of empires, free countries, and unions. At low levels of globalization ($\tau < \tau_E$), the world contains only free countries that grow in size with trade opportunities. There are no empires or unions. As globalization proceeds, the size of countries grows and so does the preference mismatch. When globalization crosses the first threshold ($\tau_E \leq \tau < \tau_U$), the preference mismatch has grown too large and the great powers prefer to form empires. The latter allows the metropolis

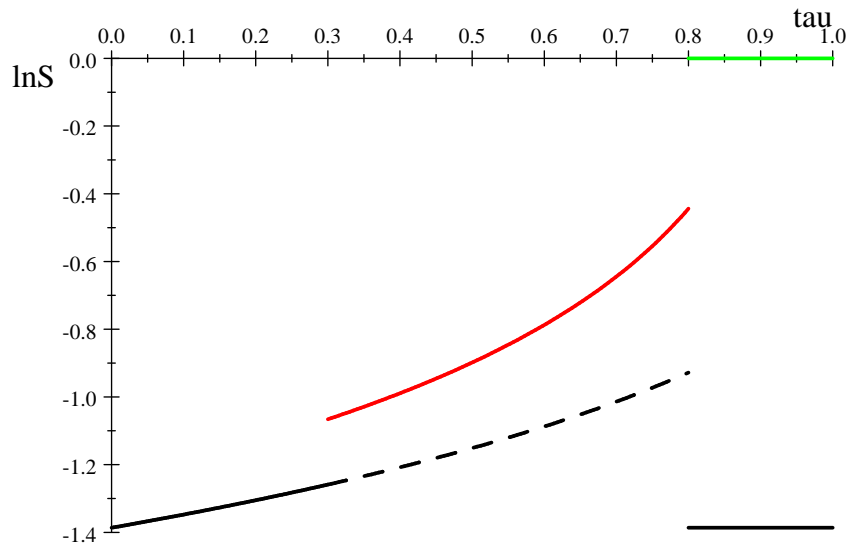


Figure 6: Countries, Empires and Unions. The figure shows how the world political structure changes with globalization (τ). The black line is the size of peaceful countries, the red line is the size of empires, the green line is the world union.

to impose their ideal type of government on the colonies. Thus, empires facilitate trade and generate economies of scale at the cost of an unbounded preference mismatch in the colonies. Empires are larger than countries and keep growing as globalization proceeds. Eventually, globalization crosses the second threshold ($\tau_U < \tau$). At this point, preference mismatch has grown too large even within empires. Empires collapse and countries revert to a smaller size. A world union is created. After this, there are no further changes in political structure.

Naturally, it is possible for some configuration of parameters that the world skips one or two of these stages. For instance, the second stage may not exist if the cost of war is high (high values of ω and μ). Or the third stage might not exist if economies of scope are strong (high values of κ). Another possibility not shown in the picture is that, before the great powers decide to abandon imperial conquest and form the world union, all localities that don't belong to empires form a smaller free union. Such a free union is never optimal if it is too small because there are too many empires (high values of π). In any case, our objective is not to provide a detailed taxonomy of cases but instead to show that our simple model is capable of helping us understand the relationship between globalization and political structure.

4 WORLD, REGIONS AND COUNTRIES

The world so far lies in an abstract space in which all localities are symmetric and equidistant. In reality, geography is much more complex. The world is made of continents and regions, and more proximate areas tend to be more integrated, both economically and politically. We can capture some of these aspects with simple modifications of our baseline model. Besides the added realism, this extension will help us rationalize why the process of formation of unions tends to be gradual and may start at the regional level.

4.1 REGIONS

Suppose that the world is divided into $r \in \mathbb{N}$ symmetric regions, each of them containing a measure $1/r$ of localities. There is a measure τ_r of goods that are traded only within the region, among $1/r$ localities, and a measure τ_w of goods that are traded with all localities in the world. As before, consumption of nontraded industries is $c_l(m, i) = e^{-\eta}$ and consumption of world-traded industries is $c_l(m, i) = e^{-b_{lm}}$. In industries that are traded only regionally, instead, there is partial specialization: $c_l(m, i) = e^{-\eta}$ if m and l do not belong to the same region and $c_l(m, i) = e^{-b_{lm}}$ otherwise. Thus, utility derived from consumption of market goods in a locality belonging to region \mathcal{R}_l is:

$$W_l^M = -(1 - \tau_w - \tau_r r) \eta - \tau_w \int_0^1 b_{lm} dm - \tau_r \int_{m \in \mathcal{R}_l} b_{lm} dm. \quad (25)$$

Notice that borders within a region have a cost of $(\tau_w + \tau_r) \beta$, while borders across regions only cost $\tau_w \beta$. This is because the volume of trade is higher within than between regions. It follows immediately that, other things equal, it is better to form countries (or unions) with localities in the same region.

Another important aspect that we want to capture is the possible difference in the cost of forming regional and world unions. To do so, we generalize our description of economies of scope in the provision of government services. So far, we have assumed that the cost of providing different economic and cultural types of government does not depend on the identity of these types. This was a natural assumption, given the symmetry of the model. However, once we introduce regions, it is equally reasonable to think that sharing government types with localities within the same region is less costly than sharing them at the world

level. Hence, we now assume:

$$\kappa = \begin{cases} \kappa_r & \text{if } g_l^E(x) = 0 \text{ for all } x \notin \mathcal{R}_l \\ \kappa_r + \kappa_w & \text{otherwise} \end{cases}. \quad (26)$$

The higher cost of sharing economic regulations at the world level may capture the higher coordination and transaction costs when providing types that are too far apart.

4.2 DIPLOMACY IN A WORLD OF REGIONS

We can now compare welfare across different political structures. As a first step, we focus on the case without empires, i.e., when the cost of war is prohibitively high. Assume for simplicity that $\sqrt{\phi/(\delta - \beta)} < 1/r$, which implies that countries are always smaller than regions. If we identify regions with real-world continents, this seems to be the empirically more relevant case, especially if we disregard colonial powers. Then, the welfare generated by a world of countries is:

$$W_N = \max_{S_N \in [0, 1/r]} \left\{ - \left(1 - \tau_w - \frac{\tau_r}{r} \right) \eta - \tau_w \beta (1 - S_N) - \tau_r \beta \left(\frac{1}{r} - S_N \right) - \delta S_N - \frac{\phi}{S_N} \right\}, \quad (27)$$

where optimal country size is still given by Equation (12) with $\tau = \tau_w + \tau_r$.

A world of regional unions delivers welfare:

$$W_R = \max_{S_U \in [0, 1/r]} \left\{ - \left(1 - \tau_w - \frac{\tau_r}{r} \right) \eta - \tau_w \beta \left(1 - \frac{1}{r} \right) - \delta S_U - \frac{\phi}{S_U} - \kappa_r \right\}, \quad (28)$$

where optimal country size is still given by Equation (14) and we have already used the result that the optimal size of a regional union is the entire region. Finally, welfare in a world union is:

$$W_U = \max_{S_U \in [0, 1/r]} \left\{ - \left(1 - \tau_w - \frac{\tau_r}{r} \right) \eta - \delta S_U - \frac{\phi}{S_U} - \kappa_r - \kappa_w \right\}, \quad (29)$$

with the same optimal country size S_U .

Once again, the equilibrium political structure is the one associated with the highest welfare. As in the benchmark model, for $\tau_r = \tau_w = 0$, the world is made of countries only. However, an increase in either τ_r or τ_w , raises welfare in all types of unions relatively more than in countries. Hence, if the cost of forming unions is sufficiently low, then globalization will change the equilibrium political structure from countries to unions.

But what kind of union will arise? Comparing W_R and W_U shows that the regional union is preferred to the world union if

$$\frac{\tau_w}{\kappa_w} < \frac{1}{\beta} \frac{r}{r-1}. \quad (30)$$

Intuitively, the regional union is preferred when the additional hurdle of a global government is high (κ_w) and when most of trade is regional, i.e., when regions are large (low r) and world tradables are scarce (low τ_w). If $\kappa_w > \beta(1 - 1/r)$, then the regional union always dominates the world union and the latter will never appear. If instead $\kappa_w < \beta(1 - 1/r)$, regional unions may still arise first, but may be followed by world unions if τ_w becomes sufficiently high. Finally, since a world union facilitates trade in all traded industries, any increase in $\tau = \tau_w + \tau_r$ raises welfare in the world union weakly more than in the regional union. It follows that regional unions may precede the world union but cannot follow it.

4.3 WAR AND CONQUEST IN A WORLD OF REGIONS

We now return to the model of war and conquest. If empires are sufficiently few in number and small in size to ensure that regions always contain free countries, then the analysis in the previous section can be extended in a straightforward way to the case with regions. There is, however, a different scenario that seems worth exploring. Suppose that all the great powers are concentrated in one region and all the localities of that region are great powers. This case corresponds to a world in which one area, which we call the “West,” has developed a superior military technology. We now briefly sketch the equilibrium political structure focusing on localities in the West.

The assumption that great powers can successfully coordinate to avoid conflict between themselves implies that the West will be evenly split into imperial metropolises and each empire will conquer colonies located in other regions. While localities prefer to form peaceful countries within a region, great powers that are geographically proximate must seek far away colonies to avoid clashing with each other. We also maintain the assumption that the world is sufficiently large that the great powers can always conquer their desired empires and that the imperial metropolis does not extend outside the entire region.

The welfare W_E of a metropolis in an empire outside of any economic union is:

$$W_E = \max_{E \in [0, 1/\tau\mu]} \left\{ - \left(1 - \tau_w - \frac{\tau_r}{r} \right) \eta - \tau_w \beta (1 - E) - \tau_r \beta \left(\frac{1}{r} - M \right) - \delta M - \frac{\phi}{E} - \omega \right\}, \quad (31)$$

with $M = \mu E$, and the preferred size of an empire is:

$$E = \sqrt{\frac{\phi}{\delta\mu - \beta(\tau_w + \tau_r\mu)}}. \quad (32)$$

Consider now unions. Note that, by definition, a regional union cannot include outside colonies (or else the cost of forming it would be $\kappa_r + \kappa_w$). It follows that if great powers decide to form a regional union, it will be of peaceful countries with corresponding welfare W_R . Moreover, under the maintained assumption that Condition (22) holds, world unions are also ruled by diplomacy, so that they generate welfare W_U . Once again, the equilibrium political structure will correspond to $W_P = \max\{W_N, W_E, W_R, W_U\}$.

As in the previous section, globalization increases the benefit of war relative to peaceful country formation. At the same time, however, globalization increases welfare in unions more than in countries and empires. Hence, countries can be replaced by empires, and both countries and empires can be replaced by unions.

Figure 7 shows how the equilibrium political structure changes with globalization, measured by $\tau = \tau_w + \tau_r$, in the most interesting four-stage scenario. At low levels of globalization, there are only countries composed of localities belonging to the same region. The size of these countries is represented by black line. As τ increases, the great powers in the West start to build large empires, conquering colonies in other regions. During these initial stages, the size of countries and empires grow with globalization. At some point, however, τ becomes so large that great powers prefer to form peaceful unions. Once the war technology is dismissed, however, political integration is more valuable at the regional level. Hence, cross-regional empires are replaced by regional unions. Finally, as globalization proceed even further, regional unions are replaced by the world union. As in the previous section, depending on the values of parameter, it is possible that the world skips some of these stages. Rather than considering all the possible cases, we next use the model to discuss various historical episodes.

5 AN INTERPRETATION OF HISTORICAL EXPERIENCE

We conclude by using our analytical results to attempt a suggestive narrative of the political evolution of modern Europe (and the world). Since the late Middle Ages, European sovereign states on average grew in size until the end of the 19th century, when this trend was dramatically reversed. For example, Kitamura and Lagerlöf (2015) show that border

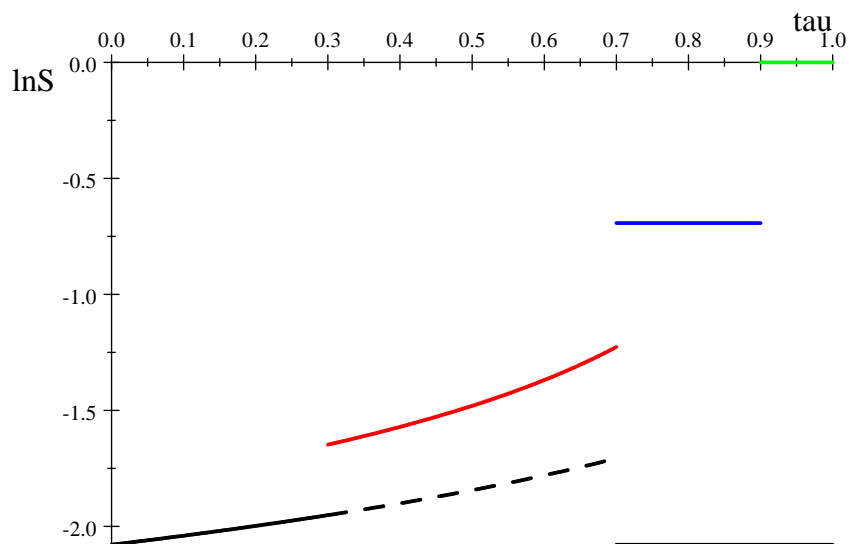


Figure 7: Countries, Empires and Regional Unions. The figure shows how the world political structure changes with globalization for the case $\tau_r = \tau_w = \tau/2$. The black line is the size of peaceful countries, the red line is the size of empires, the brown line is the regional union and the green line is the world union.

areas declined monotonically from 1500 to 1900, and then started to increase. Our model seems particularly appropriate to explain this evolution. Medieval Europe was fragmented into hundreds of small states at a time when trade was insecure and limited to few commodities. The early modern period saw important changes in both the economic and the political organization of the continent. With the Commercial Revolution, trade began to flourish and the feudal system started to be replaced by a smaller number of countries of growing size. While in 1600 there were 112 sovereign states in Europe and the Near East, at the beginning of 1800 the number had fallen to 79.

The Industrial Revolution gave trade an even more prominent role and triggered major changes in socioeconomic conditions that ultimately made the rise of the nation state possible. Trade expansion was enabled by the introduction of canals, improved roads and railways. At the same time, the high degree of political fragmentation at the time of the Congress of Vienna (1815) was followed by the unification of Germany and Italy (1871) and the further consolidation of other nation states. The economic rationale of building large internal markets was especially evident in the case of German unification, which started with the formation of a custom union (*Zollverein*) and had as a major consequence the drastic improvement in transportation infrastructure. This process of political centralization culminated at the beginning of 1900, when Europe was dominated by just 28 independent

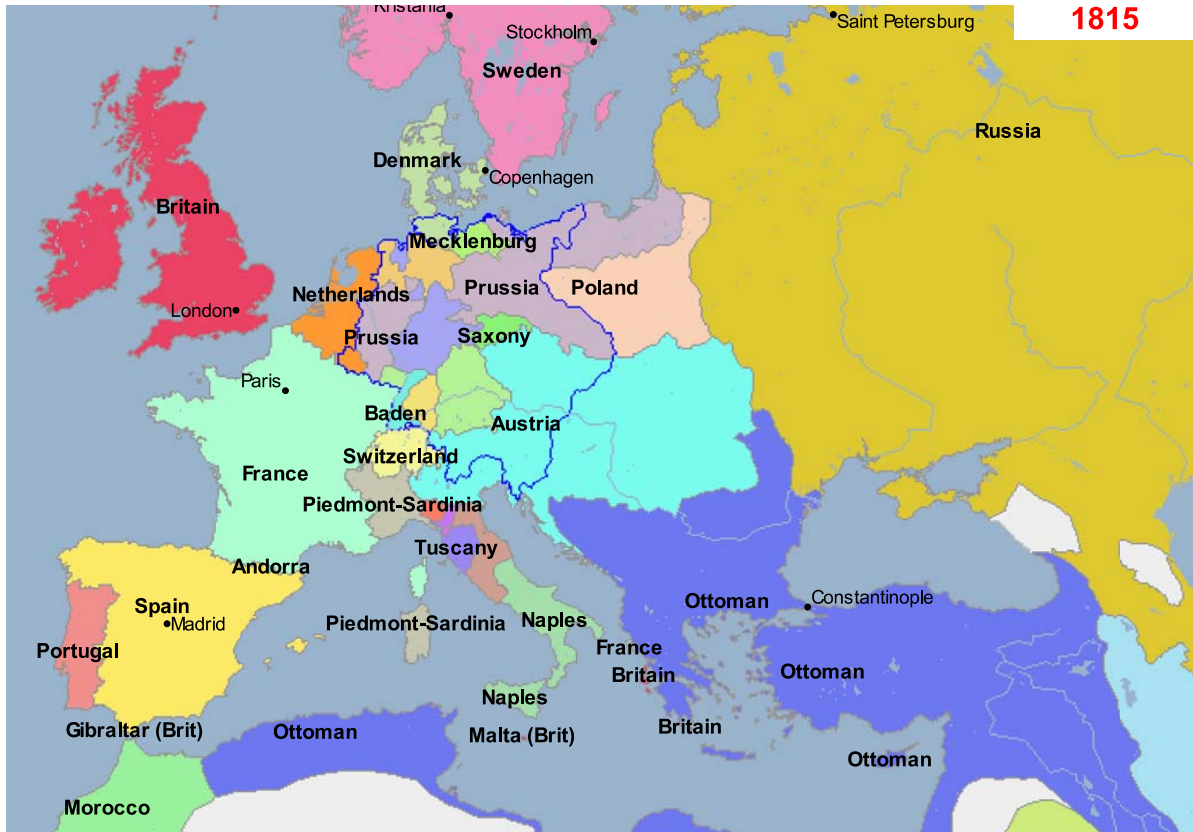


Figure 8: Europe after the Congress of Vienna (1815)

states.

Yet, the twentieth century marks a turning point. It saw the rise of international organizations both at the global level, such as the League of Nations, and at the regional level, such as the European Community. One of the key objectives of many of these supra-national organizations was precisely to promote free markets. Simultaneously, Europe entered a stage of political fragmentation, with the number of independent states growing to 58 in 2000. This pattern of an initial decline and subsequent increase in the number of countries is not confined to Europe only. For instance, the number of African countries fell from 36 in 1816 to 4 in 1914, to rise again to 51 in 2000. Similarly, in South-East Asia, this numbers changed from 37 to 4 and then 20 in the years 1816, 1914 and 2000, respectively.¹⁴ However, to better interpret the political evolutions in these regions, it is important to bring conflict into the picture.

Our model of war and conquest seems broadly consistent with the rise and fall of colonial

¹⁴The number of countries is taken from Butcher and Griffiths (2013).



Figure 9: Europe before World War I (1914)

empires. In our theory, empires are built to extract trade surplus from the colonies and disappear when the union is formed to foster free markets. According to historians and in line with this view, one of the key driving forces behind colonial expansion was the desire to secure trade and access to scarce resources in an era of revived commerce, but when mercantilist practices were common. The scarcity of land and the desire to avoid powerful rivals induced European great powers to expand by conquering territories overseas. Despite some notable setbacks, colonialism continued to grow prior to World War I and finally collapsed after World War II. The sharp decline of empires started after the creation of international agreements aimed at promoting economic cooperation. This is no coincidence. In the words of Spruyt (2005) and Rosecrance (1986), empires dissolved, often peacefully, because the gains through commerce displaced gains through territorial acquisition.¹⁵ There is also evidence that international organizations played a direct role in the process of decolonization. For example, in 1960 the UN General Assembly voted the Declaration on the Granting of

¹⁵Bonfatti (2012) also attributes the fall of empires to the growing importance of trade between industrial countries relative to trade with colonies.



Figure 10: Europe Today (2014)

Independence to Colonial Countries and Peoples.

Focusing on size, the model shares with Alesina and Spolaore (2003) the prediction that empires should be larger than democratic countries, a result which is confirmed by a quick look at Figure 2. Moreover, the model suggest that countries may form and grow for the desire to increase military might and build an empire or embark into colonial adventures. This is indeed one of the recognized reasons behind the unification of Germany. Interestingly, our theory also suggests that great powers switch to the union at higher levels of globalization than consensual countries.¹⁶ Consistently, Figure 2 shows that size started to fall earlier on for countries than for empires. Finally, the model predicts conflict to initially increase with globalization and then fall drastically after the formation unions. Data on changes of borders from Correlates of War seem supportive. The percentage of peaceful changes of borders fell from 70% in the period 1816-1900 to 62% in the years 1901-1950, and then rose to 89% in the period 1951-2008. Moreover, promoting peace by tightening economic integration was

¹⁶Formally, the emergence of empires unambiguously retards the creation of the world union. Moreover, unions of free countries may emerge earlier, with empires joining later on.

among the explicit goals of the European Union and the WTO.

The model can also be used to interpret the foundation and growth of the United States. Improvements in transportation technology and the desire to create a large internal market were important factors in its westward expansion. The abundance of land made it possible to create one of the largest countries in the world, without the need to seek far away colonies. Despite its size, the United States avoided the phase of collapse and political fragmentation by choosing an institutional system with multiple levels of government. In this light, the experience of the United States follows the main pattern predicted by our theory: the creation and expansion of the federal government, which can be interpreted as a regional union, coincided with the fragmentation and loss of political power of individual states.

After the Declaration of Independence, the borders of the original thirteen states extended up to the Mississippi river while the remaining land was occupied by French and Spanish colonies (later on part of Mexico) and by many tribes of native Americans living essentially in autarky. As the federal government acquired land and built roads westward, its territory was gradually fragmented into the fifty states. This process followed a common pattern. First, new land was annexed as large “territories;” subsequently this land was broken into new states. Federal expansion was followed not just by the creation of new states, but also by the break-up of existing ones. For example, the former Republic of Texas (an independent nation until 1846) encompassed large parts of current Oklahoma, Kansas, Colorado, Wyoming and New Mexico; similarly, Georgia, Massachusetts, North Carolina and Virginia all lost land to form new states.

Finally, our theory has been motivated by modern and contemporary political and economic events because our model of trade and globalization seems especially suited to study the period after the Industrial Revolution. Yet, some of its key implications seems consistent with ancient history as well. For instance, several historians stress the importance of trade and market size for the expansion of countries. According to Pirenne (1925, 1939), the Roman civilization was heavily dependent upon Mediterranean trade and it collapsed when trade ended with the Arab conquest.¹⁷ In his view, the cutting of major trade routes forced individual regions into self-sufficiency and this contributed to the consequent decline and fragmentation of Western civilization into the Middle Ages.

¹⁷Similarly, Friedman (1977) argues that trade increases the value of land and hence promotes territorial expansion.

6 CONCLUSION

In this paper, we have studied how the forces of globalization contribute to shape the world's political structure. Our theory has shown that the expansion of trade opportunities can help explain two salient puzzling phenomena in recent history. First, the rise and subsequent fall in the size of countries observed during the 19th and 20th century. Second, the seemingly contradictory trends towards more political integration across countries and more political fragmentation within countries in the second half of the 20th century. Our theory is broadly consistent with a variety of historical episodes. Despite its simplicity, our model has proved capable of providing a remarkably rich picture. Yet, there are several important factors that we have deliberately left aside. We now briefly mention three that seem particularly promising for future research.

First, in this paper we have focused on economic globalization as an expansion of trade opportunities. However, in reality globalization is a much more complex process that also affects cultural preferences. If cultural globalization lowers the preference for heterogeneity, it will also reduce the cost of removing borders. Hence, cultural globalization may reinforce political integration, but may also lower the value of economic unions. The equilibrium political structure may then be the result of a race between economic and cultural globalization. In a similar vein, cultural preferences may be affected by the political choices. For example, historically governments have often taken actions aimed at homogenizing their populations (Alesina and Reich 2015).

Second, we have modeled economic unions as agreements aimed at facilitating trade. Although this approach is both simple and realistic, it abstracts from another important role of unions, namely, to solve cross-border externalities associated with domestic government activity (Alesina, Angeloni and Etro 2005). Since globalization is likely to exacerbate such policy externalities (Broner and Ventura 2011; Epifani and Gancia 2009), it increases the value of forming unions. However, in this case international agreements must be properly designed to eliminate any incentive for individual countries to free ride.

Finally, our concept of locality abstracts from internal heterogeneity both along cultural and economic lines. Yet, historical experience suggests that internal conflict has played a role in many processes of country formation and break-up (Bolton and Roland 1996, 1997). It would be interesting to see how globalization also affects political structure through its effect on domestic heterogeneity and conflict.

REFERENCES

- [1] Alesina, Alberto, Ignazio Angeloni and Federico Etro (2005). “International Unions,” *American Economic Review* 95, 602-15.
- [2] Alesina, Alberto and Bryony Reich (2015). “Nation Building,” NBER Working Paper No. 18839.
- [3] Alesina, Alberto, and Enrico Spolaore (1997). “On the Number and Size of Nations,” *Quarterly Journal of Economics* 112: 1027–1056.
- [4] Alesina, Alberto, and Enrico Spolaore (2003). *The Size of Nations*. Cambridge, MA: MIT Press.
- [5] Alesina, Alberto, Enrico Spolaore and Romain Wacziarg (2000). “Economic Integration and Political Disintegration,” *American Economic Review* 90(5): 1276–1296.
- [6] Besley, Timothy, and Stephen Coate (2003). “Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Approach,” *Journal of Public Economic* 87: 2611–2637.
- [7] Boffa, Federico, Amedeo Piolatto, and Giacomo A. M. Ponzetto (2015). “Political Centralization and Government Accountability,” *Quarterly Journal of Economics*, forthcoming.
- [8] Bolton, Patrick, and Gerard Roland (1996). “Distributional Conflicts, Factor Mobility and Political Integration,” *American Economic Review* 86(2): 99–104.
- [9] Bolton, Patrick, and Gerard Roland (1997). “The Breakup of Nations: A Political Economy Analysis,” *Quarterly Journal of Economics* 112(4): 1057–1090.
- [10] Bonfatti, Roberto (2012). “Trade and the Pattern of European Imperialism, 1492-2000,” Oxford University Department of Economics Discussion Paper No. 618.
- [11] Broner, Fernando and Jaume Ventura (2011). “Globalization and Risk Sharing,” *Review of Economic Studies* 78(1): 49-82.
- [12] Butcher, Charles, and Ryan Griffiths (2013). “Introducing the International System(s) Dataset (ISD), 1816-2011.” *International Interactions* 39(5): 748–768.

- [13] Casella, Alessandra (2001). “The Role of Market Size in the Formation of Jurisdictions,” *Review of Economic Studies* 68: 83–108.
- [14] Casella, Alessandra, and Feinstein, Jonathan S. (2002). “Public Goods in Trade: On the Formation of Markets and Jurisdictions,” *International Economic Review* 43(2): 437–462.
- [15] Epifani Paolo and Gino Gancia (2009). “Openness, Government Size and the Terms of Trade,” *Review of Economic Studies* 76, 629–668.
- [16] Friedman, David D. (1977). “A Theory of the Size and Shape of Nations,” *Journal of Political Economy* 86(1): 59–77.
- [17] Hugot, Jules (2015). “Trade Costs and the Two Globalizations: 1827-2012.” Working paper, Pontificia Universidad Javeriana.
- [18] Kitamura, Shuhei, and Nils-Petter Lagerlöf (2015). “Natural Borders.” IIES Seminar Paper No. 773.
- [19] Lockwood, Benjamin (2002). “Distributive Politics and the Costs of Centralization,” *Review of Economic Studies* 69(2): 313–337.
- [20] Maddison, Angus (1995). *Monitoring the world economy, 1820-1992*. Paris: OECD.
- [21] Maddison, Angus (2001). *The World Economy: A Millennial Perspective*. Paris: OECD.
- [22] Martin, Philippe, Thierry Mayer and Mathias Thoenig (2008). “Make Trade Not War?” *Review of Economic Studies* 75(3): 865–900.
- [23] Oates, Wallace E. (1972). *Fiscal Federalism*. New York, NY: Harcourt Brace Jovanovich.
- [24] Pascali, Luigi (2014) “The Wind of Change: Maritime Technology, Trade and Economic Development.” CAGE Working Paper No. 195.
- [25] Pirenne, Henri (1925). *Medieval Cities: Their Origins and the Revival of Trade*. Princeton, NJ: Princeton University Press.
- [26] Pirenne, Henri (1939). *Mohammed and Charlemagne*. New York, NY: Norton.
- [27] Rosecrance, Richard (1986). *The Rise of the Trading State: Commerce and Conquest in the Modern World*. New York, NY: Basic Books.

- [28] Ruta, Michele (2005). “Economic Theories of Political (Dis)integration,” *Journal of Economic Surveys* 19: 1–21.
- [29] Singer, David J., and Melvin Small (1966). “The Composition and Status Ordering of the International System: 1815-1940”, *World Politics* 18: 236–282.
- [30] Spruyt, Hendrik (2005). *Ending Empire: Contested Sovereignty and Territorial Partition*. Ithaca, NY: Cornell University Press.

A APPENDIX

A.1 DATA SOURCES

The trade share reported in Figure 1 is merchandise exports as percent of GDP in 1990 prices, from Maddison (1995, 2001). Maddison provides trade data for selected countries in the years 1820, 1870, 1913, 1929, 1950, 1973 and 1998. To avoid compositional effects, we report the value of merchandise export as a share of GDP for the set of countries with data for all the years (Austria, Belgium, France, Italy, Spain, Switzerland, the United Kingdom and the United States). The trade share computed using the data for all available countries in every year is very similar to the one displayed in Figure 1.

The number of countries is reported for the same years. Data on the number of countries in the twentieth century is not very controversial. For the nineteenth century, however, some leading conventions grossly underestimate the number of countries. For example, the “International System” developed by Singer and Small (1966) and adopted in the Correlates of War project or in the Cross-National Time-Series Data Archive, only includes countries with international recognition. In particular, prior to 1920, the criteria to be recognized as an independent country were to have population greater than 500,000 and have had diplomatic missions at or above the rank of charge d’affaires with Britain and France. Clearly, this definition is too strict for our purposes, which require the identification of even relatively small political units living in economic and political autarky. We follow Butcher and Griffiths (2013) who recognize the problem and offer alternative sets of criteria to identify the number of countries between 1816 and 2011.

The number of WTO members is from the WTO’s website.

The size of countries and empires displayed in Figure 2 is from the Cross-National Time-Series Data Archive (CNTS). It provides data on contiguous territorial area in thousand square miles for all countries existing in a given year according to the International State System. In a few instances, missing data have been imputed cross-checking major territorial changes from other sources (China and Persia before 1860). Area of empire is provided for a consistent sample of 13 countries: Austria-Hungary (later Austria), Belgium, France, Germany (Prussia), Italy (Sardinia), Japan, Netherlands, Portugal, Russia, Spain, Turkey (Ottoman Empire), United Kingdom, and United States. For these countries, empire area includes “overseas” territories (i.e., colonies). Data for the two modern wartime periods, 1914-1918 and 1940-1945 (1938-1954 for Empires) are missing.

European political maps were drawn using the online software GeaCron (<http://www.>

geacron.com). This software provides a geo-temporal database that can be used to draw geopolitical maps of any region in the world, in any given historical time period. The number of independent states in Europe and Near East reported in Section 5 is taken from Euratlas-Nüssli (<http://www.euratlas.com>).

A.2 COMPUTING EQUILIBRIUM CONSUMPTIONS

Locality l maximizes

$$W_l^M = \int_0^1 \int_0^1 \ln c_l(m, i) dm di \quad (\text{A1})$$

subject to the following budget constraint:

$$\int_0^1 \int_0^1 p_l(m, i) [c_l(m, i) - q_l(m, i)] dm di \leq 0, \quad (\text{A2})$$

where $q_l(m, i)$ and $p_l(m, i)$ are the production and price of input m of industry i in locality l . The productions $q_l(m, i)$ must be consistent with available technology as described in the text. Since individuals are atomistic, they take prices as given in their maximization problems.

We claim now that equilibrium prices are given as follows:

$$p_l(m, i) = \begin{cases} b_{lm} & \text{if } i \in [0, \tau] \\ \eta & \text{if } i \in (\tau, 1] \end{cases}. \quad (\text{A3})$$

To prove this claim, note first that all localities spend the same on each input: di . Thus, the world share of spending on each input is also di . Let world income be Y . Then, the spending on each input is Ydi . We next examine productions. Consider first nontraded industries. Our assumptions on preferences and technology ensure that all localities use the specialization technology and produce $\eta^{-1}di$ units of each nontraded input. Thus, the value of the production of each nontraded input is $(\eta/\eta)di = di$. Consider next traded industries. The assumption that $b_{lm} < \eta$ ensures that all localities prefer to use the technology based on specialization. Thus, the value of the production of each traded input is $1 \times di = di$. Since world income is $Y = 1$, spending in each input is also di . This proves our claim.¹⁸

¹⁸It is straightforward to show that this equilibrium is unique. First, rule out variation in the prices of traded inputs since this would generate excess demand (supply) of cheap (expensive) varieties. Second, rule out that the relative prices of traded and nontraded varieties be above (below) b_{lm}/η since this would lead to an excess demand (supply) of nontraded inputs.

With these prices at hand, we can compute the equilibrium productions and consumptions described in the text.

A.3 DEALING WITH INTEGER CONSTRAINTS

A world without unions consists of $N^* \in \mathbb{N}$ countries of measure $S_n > 0$ each, such that $\sum_{n=1}^{N^*} S_n = 1$ and utilitarian social welfare is $W = \sum_{n=1}^{N^*} S_n W_n(S_n)$.

The welfare function

$$W_n(S_n) = -(1 - \tau)\eta - \tau\beta(1 - S_n) - \delta S_n - \frac{\phi}{S_n} \quad (\text{A4})$$

is concave in S_n and increasing at $S_n = 0$. For all $\tau < (\delta - \phi)/\beta$ it has a unique maximum at $S_n = S_N$.

Pareto efficiency then requires that either $S_n \leq S_N$ for all n or $S_n \geq S_N$ for all n , or else some localities could leave a country with excessive size $S_n > S_N$ and join another with insufficient size $S_n < S_N$, raising the welfare of every locality in both countries.

Utilitarian welfare maximization requires all countries to have the same size. If $S_m > S_n > S_N$ then transferring the marginal locality from m to n not only raises its welfare, but it also raises the welfare of S_m localities by more than it lowers the welfare of $S_n < S_m$ localities. Likewise if $S_m < S_n < S_N$.

Therefore, once integer constraints are taken into account, the utilitarian welfare optimum without unions is a partition of the world into a number

$$N^* = \arg \max_{N \in \mathbb{N}} \left\{ \frac{\tau\beta}{N} - \frac{\delta}{N} - \phi N \right\} \quad (\text{A5})$$

of identical countries. The maximand has strictly decreasing differences in (N, τ) because for any $\tau_H > \tau_L$ and $N_H > N_L$,

$$\begin{aligned} W(N_H, \tau_H) - W(N_H, \tau_L) &= (\tau_H - \tau_L) \frac{\beta}{N_H} \\ &< W(N_L, \tau_H) - W(N_L, \tau_L) = (\tau_H - \tau_L) \frac{\beta}{N_L}. \end{aligned} \quad (\text{A6})$$

Thus, the welfare-maximizing number of countries N_C^* is decreasing in τ in the sense of monotone comparative statics. It is likewise decreasing in β , and increasing in δ and ϕ .

By the same reasoning, the utilitarian welfare optimum with unions is a partition of the

world into atomistic countries joined into a number

$$N_U^* = \arg \max_{N \in \mathbb{N}} \left\{ \frac{\tau\beta}{N} - \phi N \right\} \quad (\text{A7})$$

of identical unions.