School of Arts and Social Sciences
CITY UNIVERSITY LONDON

Department of Economics

Schooling, Nation Building, and Industrialization: a Gellnerian Approach

Esther Hauk
Institut d'Anàlisi Econòmica (IAE-CSIC) and Barcelona Graduate School

Javier Ortega
City University London, CEP (LSE), CReAM and IZA

Department of Economics
Discussion Paper Series
No. 15/06

Schooling, Nation Building, and Industrialization: a Gellnerian Approach*

Esther Hauk[†] Javier Ortega[‡]

July 2015

Abstract

We model a two-region country where value is created through bilateral production between masses and elites (bourgeois and landowners). Industrialization requires the elites to finance schools and the masses to attend them. Schooling raises productivity, particularly for matches between masses and bourgeois. At the same time, only country-wide education ("unified schooling") renders the masses mobile across regions. Alternatively, schools can be implemented in one region alone ("regional education") or the regionally dominant group can choose to implement schooling in its own region but refuse to share the costs/proceeds within the wider country-level group ("secession"). We show that schools are more likely to be set-up when the bourgeoisie dominates, but that this is not necesarily socially efficient. Unified schooling is always chosen if the identity of the dominant elite at the regional and country level is the same and/or the industrialization shock is sufficiently high. If instead the bourgeoisie is dominant in one region and landowners are dominant countrywise, the bourgeoisie of that region may promote the secession of the region, and this can be socially efficient. The model is shown to be consistent with evidence for 19th century France and Spain.

JEL: D02, I2, N00, O14.

Key words: Nation-building, education, industrialization.

*We thank Ramon Caminal, Paolo Masella, Yona Rubinstein and seminar participants at the CEP (LSE), Royal Economic Society, ESSLE-CEPR, City University London, ENSAI (Rennes), Workshop on Culture and Institutions (Valencia), and Birkbeck for useful comments and suggestions. Hauk thanks the LSE for its hospitality and acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2011-0075) and through CICYT project number ECO2012-37065 and from the government of Catalonia.

†Institut d'Anàlisi Econòmica (IAE-CSIC) and Barcelona Graduate School of Economics, Campus UAB, Bellaterra (Barcelona); email: esther.hauk@iae.csic.es.

‡City University London, CEP (LSE), CReAM, and IZA, Northampton Square, London EC1V 0HB, e-mail: javier.ortega.1@city.ac.uk

1 Introduction

Political scientists, historians, sociologists and anthropologists have extensively discussed the issue of the historical genesis of nations and nationalism (see e.g. Smith, 2000, for a summary of the debate). While "perennialists" argue that national identities have existed for a long period of time (see e.g. Armstrong, 1982, or Hastings, 1997), "modernists" situate the birth of nations and nationalism during industrialization.

In particular, Gellner (1964, 1983) has been very influential in arguing that both Nations and Nationalism result from the implementation of mass educational systems to get workers ready for industrialization. As stated by Breuilly (2006, p. xxxiv), "Gellner insisted that industrialization required or entailed cultural homogenization based on literacy in a standardized vernacular language conveyed by means of state supported mass education". According to Gellner, industrialization requires a diffuse, universal culture, linking the inhabitants of a territory to the state. Because workers, through schooling, acquire a common national identity that enables them to communicate with each other, they also become mobile. In addition, as mass education is expensive, Gellner (1983) argues that the minimum size for a viable modern political unit is determined by the ability to finance such an educational system. More recently, Breuilly (1993) has criticized Gellner's theory and other theories of nationalism because they failed to stress that nationalism is about power and state control, and has argued that "the central task is to relate nationalism to the objectives of obtaining and using state power" (Breuilly, 1993, p. 1). In addition, Roeder (2007) and Kroneberg and Wimmer (2012) argue that nation building should be understood as resulting from the interaction between central and peripheral elites.[1]

We contribute to the literature by developing a theoretical model that explicitly takes into account the interaction of social groups holding power and relates nation building, schooling and industrialization *à la* Gellner.

To this purpose, we model a two-region economy populated by masses and by two elite groups (landowners and bourgeoisie, as in Galor, Moav and Vollrath, 2009) who live for two periods. Regions are heterogeneous in the size of their bourgeoisie. Political power is in the hands of one of the elite groups, referred to as the "dominant group", which is not necessarily the same at the regional and at the country level. Value is created through bilateral production between the members of the elites and the members of the masses. Initially, the country is a rural society, and production takes place only within each region.

The economy is hit by a productivity shock representing an industrialization opportunity which can raise the productivity of the masses but only if they attend school.[2] In that case

[1] The importance of the power interaction among groups in the genesis of institutions has been extensively studied in the literature, see e.g. Acemoglu and Robinson (2001).

[2] There is a debate on whether industrialization caused mass schooling or the other way round. What matters for our model is that these phenomena go hand in hand. Becker, Hornung and Woesmann (2011) reveals the importance of formal education for the technological catch-up of Prussia. Galor and Moav (2006) gives historical evidence for the industrial base for education reforms in the 19th century and reveals the importance of schooling for at least the second phase of the industrial revolution. At the same time, Allen (2003) argues that the impact of literacy on growth was limited and Squicciarini and Voigtlaender (2014) shows that knowledge of the elites (and not literacy) predicts growth in France between 1750 and 1850. For an alternative hypothesis for the

productivity is raised to a larger extent in the matches with bourgeois than with landowners.[3] In addition, schooling is used to create a common identity.[4] The set-up of the schooling system can only be financed by the elites, but mass members decide whether to attend school, which forces them to forego production in the first period.

The politically-dominant elite group decides if and how to implement schooling and how the costs of schooling are shared within the elite. In particular, the politically-dominant country-level elite can choose to implement schooling in one region only ("regional education") in which case only within-region production is possible. Alternatively, it can choose to implement schools in both regions ("unified schooling"), which creates a common national identity and makes it possible for the masses of one region to produce with the other region's bourgeoisie. Finally, we consider the possibility that the dominant *region*-level elite implements schooling in its own region but refuses to share the associated costs and benefits within the wider country-level group ("secession").

Under all three systems, equilibrium education is shown to be weakly higher when the bourgeoise dominates, which stems from the higher payoffs of bourgeois relative to landowners.

However, whenever the dominant group is the same at the country and regional level, the identity of the dominant group does not matter for the choice of the type of educational system. Indeed, in that case, unified schooling is always chosen given its technological advantage. Specifically, a dominant bourgeoisie prefers this system because it can directly benefit from the increase in the pool of matches, while dominant landowners also favor it because the bourgeois are willing to pay a larger share of the schooling cost under this system.

However, despite this technological advantage, unified schooling can still be dominated by secession if the dominant elite is not the same regionally and countrywide. In particular, if the bourgeoisie is regionally-dominant and countrywide-dominated, the size of the cake is larger for them under unified schooling but at the same time landowners can impose a large share of the costs on them. In that case, the bourgeoisie chooses secession for intermediate values of the shock: if the shock is small enough, no schooling is implemented under secession - but unified schooling might be implemented. In the other extreme, if the shock is large, the size-of-the-cake effect under unified schooling always dominates. We also show that regionally-dominant and countrywide-dominated landowners always prefer secession to unified schooling when schooling under secession is possible, since they can save on educational costs and do not benefit from the regional mobility of the masses.

As for welfare, unified schooling leads to the underprovision of education whenever the gains from setting up schools for the dominant group are small relative to the gains for the masses, and particularly so when landowners are dominant, as they benefit less from education than the bourgeois. More interestingly, overeducation can arise if the bourgeoisie is dominant, as this

implementation of mass education systems based on military rivalry see Aghion, Jaravel, Persson, and Rouzet (2013).

[3]The same hypothesis is made in Galor et al. (2009). Empirically, Lindert (2004) refers to examples of resistance of landlords to education in 19th century England and Germany, and Ager (2013) shows that counties with richer planters before the Civil War invested less in human capital and were less productive in the 20th century.

[4]For a formal model of schooling as an instrument for language uniformization, see Ortega and Tangerås (2008).

group chooses in some cases to fully finance education even if this makes the landowners worse-off. Across systems, a social planner always prefers unified schooling over secession whenever implementing education is socially optimal. Hence only unified schooling or no schooling at all can be first best. If instead the central planner can choose the education system but investment in schooling remains in the hands of the elites, secession can be a second best. Specifically, this arises when a regionally-dominant bourgeoisie implements education under secession while the country-wide dominant landowners choose not to implement schooling under the unified system. At the same time, we show that landowner-driven secessions always lower welfare.

We also discuss other forms of heterogeneity across regions and their effects on nation building and secession. Our results are robust to differences in sizes across the landowners and masses. However, if productivity shocks are unequally distributed across regions - a case that seems to be historically relevant - secession becomes more likely as it avoids costly transfers from the more advanced region to the less advanced region.

Finally, we show that our model can be used to interpret the divergent evolution of France and Spain in the 19th century. Indeed, despite their common features in terms of income levels and language heterogeneity at the beginning of the 19th century, France was successful in its joint nation building/industrialization process through the implementation of a large investment in education. Instead, both industrialization and nation-building remained weak in Spain, and peripheral nationalisms developed in Catalonia and the Basque Country. As predicted by our model, the divergent evolution of these two countries could be related to the different balance of power between landowners and bourgeois at the regional and country level: while in France the bourgeoisie was dominant both in the industrializing regions and at the country level, in Spain the Catalan bourgeoisie was unable to have a lot of influence in Spanish politics due to the dominance of the landowning elites at the country level.

This paper relates to a growing literature that uses modelling or econometric techniques to study the origin of nations or nation-states. Specifically, Darden and Grzymala-Busse (2006), Aspachs-Bracons *et al.* (2008) and Clots-Figueras and Masella (2013) underline the importance of education for nation-building.[5] Alternative mechanisms proposed in the literature as driving forces for nation-building include the consolidation of a previously existing "segment-state" (Roeder, 2007), political centralization prior to modernization (Kroneberg and Wimmer, 2012), or the homogenization of preferences on public goods (Alesina and Reich, 2013). Empirically, Wimmer and Feinstein (2010) argues that the origin of nation-states lies on local and regional factors.

The remainder of the paper is organized as follows. In section 2 we develop the basic model and describe when regional and unified schooling are implementable. In turn, these two systems are compared in Section 3. After introducing secession as a possible outcome in Section 4, Section 5 studies when secession will be chosen over unified schooling. Next, we study welfare

[5] Building on Darden and Grzymala-Busse (2006)'s theory that nationalism often results from the identity transmitted during the *first* implementation of mass schooling, Balcells (2013) argues that French Catalonia ended up a with a strong French identity because its inhabitants were first educated at the end of the 19th century as French patriots. Instead, the lack of investment in education by the Spanish state during the same period meant that the inhabitants of Spanish Catalonia were educated later – from the beginning of the 20th century- at a time in which Catalan nationalism was developing and was taking initiatives for the education of masses.

(Section 6) and extend the model to alternative forms of heterogeneity (Section 7). Finally, in section 8 we confront the predictions of our model with the cases of 19th century France and Spain. Section 9 concludes. Most proofs are relegated to the appendix.

2 The Model

We study a country with two regions $i = 1, 2$. In each region, there are three social groups, namely the masses $M = M_1 + M_2$ and the elite which is split into the landowners $N = N_1 + N_2$ and the bourgeoisie $B = B_1 + B_2$. Political power is in the hands of one of the elite groups, which is referred to as the "dominant" group. The dominant group holds power for historical reasons and is not necessarily the majority elite group. Moreover, while there is one dominant group at the country level, this group is not necessarily dominant in both regions. Let $M > N + B$. We normalize the total size of the elite in the country to $N + B = 1$. For simplicity, we assume that both landowners and masses are equally distributed across regions, i.e. $N_1 = N_2 = \frac{N}{2}$ and $M_1 = M_2 = \frac{M}{2}$. Instead, one region is characterized by a larger bourgeoisie than the other, and this region is assumed to be region 1, without loss of generality (i.e., $B_1 > B_2$).

Value is created through bilateral production between members of the elites and members of the masses. Initially, the country is a "rural" society. Production takes place only within each region and the surplus from each match is normalized to 1. The bargaining power of the masses is given by β, which simply implies in our framework that a member of the masses who is matched to a member of the elite keeps β of the surplus generated from the match.

This rural society is now hit by a productivity shock representing the industrial revolution. If the new technology is implemented, the match productivity in the agrarian sector (landowner-masses) increases to $1 + \sigma$ while the match productivity in the industrial sector representing a match between a bourgeois and the masses increases to $1 + \mu\sigma$ where $\mu > 1$. However, the increase in productivity only occurs if the member of the masses attends school. Otherwise, the productivity of the match remains equal to 1. In other words, the implementation of the new technology requires schooling of the masses.

The set-up of a schooling system can only be financed by the elites, but the masses decide whether or not to attend school.

There are two periods in our model: in the first period, the productivity shock is observed and the schooling decision is made. If schooling is implemented, production takes only place in the second period. If schooling is not implemented, production takes place in both periods but the match productivity stays equal to one. All agents have a discount factor of δ.

2.1 Payoffs if schools are not implemented

Let Ψ_j $(j = B, N, M_i)$ denote the payoff of a member of group j when schooling is not implemented. In this case, any member of the elite produces an output of 1 with each of the $M/2$ members of the masses living in his region, and gets a proportion $1 - \beta$ of the output. As a result, the payoff of a landowner is the same as that of a bourgeois and is given by

$$\Psi_N = \Psi_B = (1 - \beta)(1 + \delta)\frac{M}{2}. \tag{1}$$

For a member of the masses in region i, the pay-off is:

$$\Psi_{M_i} = (1+\delta)\beta\left(\frac{N}{2} + B_i\right) \tag{2}$$

i.e. the member of the masses receives β when producing with the $N/2$ landowners and B_i bourgeois in region i.

2.2 Schools

The dominant group chooses whether or not schooling is implemented and how to split the schooling costs among the elite. We assume that the dominant group cannot force the dominated group to pay for schooling if with this payment the dominated group would be made worse-off than under no schooling. This implies that the maximum schooling costs that can be imposed on the dominated elite group leave this group indifferent between the implementation of schooling and the absence of schools.

We also assume that each of the elite groups acts as a single group at the country level, i.e. each group equally shares across regions the benefits from production and the costs from schooling.

Schools can be implemented either in both regions, or in one region only. The implementation of schools in both regions creates a common identity across regions, which enables the masses of each region to produce with the bourgeois from both regions. This is referred to as nation building or a "unified" schooling system, and denoted by U. Instead, if schooling is implemented in one region only, no common identity is created, and thus the masses of each region can only produce with the bourgeois of the same region. This is referred to as a "regional" schooling system, and denoted by R_i $(i = 1, 2)$. In both cases, the masses can only produce with the landowners of their region of origin.

2.2.1 Payoffs from schooling

Let Π_j^k denote the payoffs from schooling for group $j = B, N, M_i$ under organizational system $k = U, R_i$. Similarly, denote by I_e^k the cost of setting up schooling system k for a member of the elite group $e = N, B$. We can next calculate the benefits from schooling for each group under the different systems.

When attending school in a unified system, any member of the masses foregoes production in the first period and appropriates in the second period (discounted by δ) a fraction β of the amount $1+\sigma$ produced with each of the $N/2$ landowners in his region and the same fraction of the amount $1+\mu\sigma$ produced with each of the B bourgeois in the country:

$$\Pi_{M_i}^U = \beta\delta\left((1+\sigma)\frac{N}{2} + (1+\mu\sigma)B\right) \qquad i = 1, 2. \tag{3}$$

Any bourgeois pays I_B^U schooling set-up costs, and appropriates a fraction $1-\beta$ of the amount $1+\mu\sigma$ produced with the M members of the mass in period 2, i.e.,

$$\Pi_B^U = -I_B^U + (1-\beta)(1+\mu\sigma)\delta M. \tag{4}$$

The landowner's payoff depends on its own investment I_N^U and is associated to a lower match productivity $(1+\sigma)$ and to a smaller pool of mass members than for the bourgeois, namely the $M/2$ mass members living in the landowner's region:

$$\Pi_N^U = -I_N^U + (1-\beta)(1+\sigma)\delta\frac{M}{2}. \tag{5}$$

Under region-i schooling, the payoff of any member of the masses in region i is

$$\Pi_{M_i}^{R_i} = \beta\delta\left((1+\sigma)\frac{N}{2} + (1+\mu\sigma)B_i\right) \qquad i = 1, 2 \tag{6}$$

where the only difference with (3) is that only within region-i production is possible.

In turn, each of the B_i region-i bourgeois gets $(1-\beta)(1+\mu\sigma)$ in the second period when producing with the $M/2$ educated region-i mass members, while each of the B_{-i} bourgeois in region $-i$ gets $(1+\delta)$ with the $M/2$ uneducated masses of region $-i$. Then, given cross-subsidization across regions within the countrywide bourgeoisie, the payoff of a bourgeois is given by the weighted average of these two terms minus the school set-up cost $I_B^{R_i}$, i.e.

$$\Pi_B^{R_i} = -I_B^{R_i} + (1-\beta)\left(\delta(1+\mu\sigma)B_i + (1+\delta)B_{-i}\right)\frac{M}{2B} \quad \text{for } i = 1, 2. \tag{7}$$

Finally, each of the $N/2$ region-i landowners gets $\delta(1-\beta)(1+\sigma)$ when producing with the $M/2$ educated masses of that region, while each of the $N/2$ landowners in region $-i$ gets $(1+\delta)(1-\beta)$ when producing with the $M/2$ uneducated masses of region $-i$, and thus the payoff of a landowner is:

$$\Pi_N^{R_i} = -I_N^{R_i} + (1-\beta)\left(\delta\sigma + 1 + 2\delta\right)\frac{M}{4} \qquad \text{for } i = 1, 2. \tag{8}$$

2.3 Education thresholds for the elites

In this subsection we study the minimum size of the productivity shock that makes the elite willing to provide schooling under the assumption that the masses attend school when schools are built.[6]

The minimum productivity shock that makes the elite indifferent between implementing unified schooling or not is such that $\Psi_e = \Pi_e^U$ with $e = N, B$. From (1), (4), and (5), the thresholds for the bourgeoisie and the landowners are:

$$\sigma_B^U = \frac{I_B^U + (1-\beta)(1-\delta)\frac{M}{2}}{\mu M(1-\beta)\delta} \tag{9}$$

$$\sigma_N^U = \frac{2I_N^U + (1-\beta)M}{(1-\beta)\delta M}. \tag{10}$$

[6] If the masses have the choice of whether or not to get schooled, we will additionally get a minimum productivity shock that makes the masses willing to get schooled. In this case, schooling is implemented only if the productivity shock lies above the maximum of the minimum thresholds by the masses and the elites.

Similarly, from (1), (7), and (8), the thresholds under region-i schooling are

$$\sigma_B^{R_i} = \frac{\frac{2B}{B_i} I_B^{R_i} + (1-\beta)M}{(1-\beta)\delta\mu M} \qquad \text{for } i = 1, 2 \tag{11}$$

$$\sigma_N^{R_i} = \frac{4I_N^{R_i} + (1-\beta)M}{\delta(1-\beta)M} \qquad \text{for } i = 1, 2. \tag{12}$$

All these thresholds depend on how much the elite has to pay for setting up the schools. We assume the cost of each schooling system to be proportional to the number of students attending schools and for expositional purposes set marginal schooling costs equal to 1.

The dominant elite group e determines how the costs of education are split within the elite under education system k. However, the dominant group e cannot force the dominated group $-e$ to pay for education if this payment makes the dominated group worse-off than under no schooling. In other words, the dominant group extracts from the dominated their maximum willingness to pay for education. When the dominant group chooses to implement schooling, the following cases are possible:

1. Education is sufficiently beneficial for the dominated to be willing to pay the entire cost of education. In that case, the dominant group gets education for free.
2. The dominant group has to cofinance education and pay $\widetilde{I_e^k}$ after forcing the dominated to pay the maximum acceptable amount $\overline{I_{-e}^k}$, leaving the dominated indifferent between schooling and no schooling.
3. The dominated are not willing to pay anything for education but the dominant group is better-off with education even if it pays the full cost.

Which cases will result clearly depends on the productivity shocks. The higher these shocks, the higher the potential benefits from schooling and the higher the potential willingness to pay for schooling by the dominated group. The following cutoffs will be relevant for the analysis:

Notation 1 *We denote by*

- $\widehat{\sigma}_e^k$ *the minimum productivity shock making elite group e willing to pay the entire cost of schooling.*
- $\widetilde{\sigma}_e^k$ *the minimum productivity shock making elite group e willing to cofinance education paying* $\widetilde{I_e^k}$ *when group* $-e$ *is paying its maximum willingness* $\overline{I_{-e}^k}$.
- $\underline{\sigma_e^k}$ *the minimum productivity shock making elite group e willing to implement education without paying.*

The exact values for these shocks and payments under the different educational systems can be found in Table 1 in Appendix A. Lemma 1 shows that two different rankings of the thresholds are possible depending on the attractiveness of schooling for the bourgeoisie relative to the landowners:

Lemma 1 *For $k = U, R_i$*

1. $\underline{\sigma}_B^k < \underline{\sigma}_N^k < \widetilde{\sigma}_B^k = \widetilde{\sigma}_N^k < \min\left[\widehat{\sigma}_B^k,\ \widehat{\sigma}_N^k\right]$ *if* $2 > H^k$

2. $\underline{\sigma}_B^k < \widehat{\sigma}_B^k < \underline{\sigma}_N^k < \widehat{\sigma}_N^k$ *if* $2 < H^k$

where H^k is given by

$$H^U = (1-\beta)B(2\mu - 1 + \delta) \tag{13}$$
$$H^{R_i} = 2(1-\beta)(\mu - 1)B_i \tag{14}$$

Proof. By simple algebra. ■

For a given investment in education, the gain from schooling for the bourgeoisie is larger than for the landowners because the bourgeoisie experiences a larger productivity increase than landowners and because it is the only group that might gain production partners with schooling. This explains why $\underline{\sigma}_B^k < \underline{\sigma}_N^k$ always holds.

The attractiveness of schooling for the bourgeoisie relative to the landowners is particularly high when (i) μ is very high, i.e. the bourgeoisie has a big productivity advantage over landowners, (ii) the agents discount the future to a small extent, as the future gains for schooling are higher for the bourgeoisie than for the landowners, and (iii) the size of the bourgeoisie is large, as the per capita burden from education for a bourgeois is then reduced. For this reason, when $H^k > 2$ is satisfied, the thresholds of the landowners are systematically larger than the thresholds of the bourgeoisie, and, in particular, $\widehat{\sigma}_B^k < \underline{\sigma}_N^k$ holds, i.e. there are situations (specifically, for $\widehat{\sigma}_B^k < \sigma < \underline{\sigma}_N^k$) in which the bourgeoisie is willing to set-up schools bearing the full cost while schooling for free is still not beneficial for landowners.[7] Instead, for $H^k < 2$, the attractiveness of education is more similar for both groups, and $\widehat{\sigma}_B^k > \underline{\sigma}_N^k$. In this case, the bourgeoisie's threshold for full education financing $\widehat{\sigma}_B^k$ might be bigger than the threshold for landowners $\widehat{\sigma}_N^k$ despite the extra gains from schooling for the bourgeoisie. This happens in particular if the bourgeoisie is small relative to the landowners, as in that case the *per bourgeois* cost of education is high.

Within groups, the payoff from schooling for a given elite group in a given schooling system k is decreasing in the amount paid by the group.

2.4 Provision of education by the elite

We are now in a position to represent the decision on education provision by the elite under the assumption that the masses have to follow suit[8] in a given organizational form k.[9]

[7] In this case the cutoffs for co-financing schooling are irrelevant and therefore do not appear in the second case of Lemma 1.

[8] Whether or not the masses want to follow suit will be analyzed in Subsection 2.5.

[9] We will see later on that this analysis also applies to secession (Section 4)

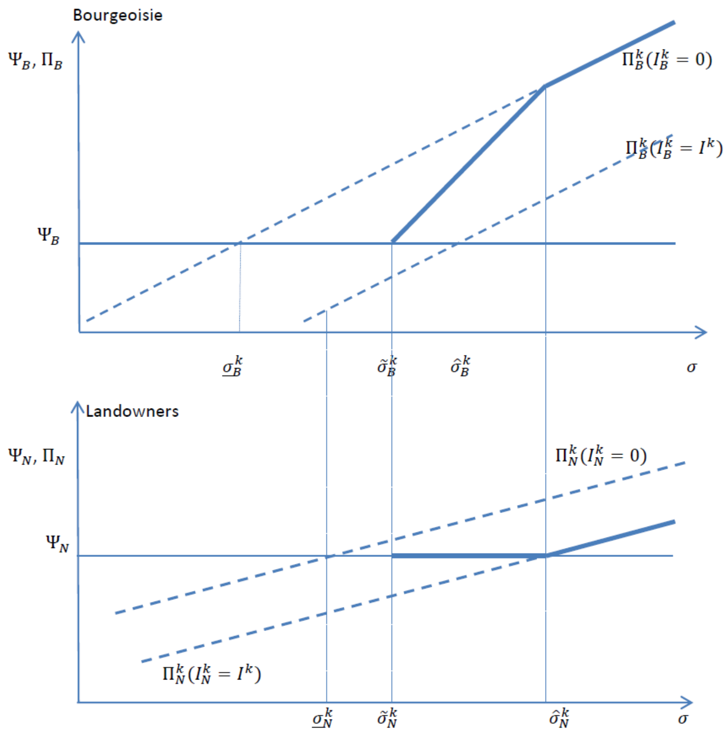


Fig. 1: Bourgeoisie dominant and $H^k < 2$

2.4.1 Bourgeoisie dominant

Figure 1 represents the decision on education provision by the elites when the bourgeoisie is dominant and $H^k < 2$. For $\sigma > \widehat{\sigma}_N^k$ the landowners are willing to pay the full cost of education, and thus the bourgeoisie puts the full burden on them. For $\widetilde{\sigma_N^k} = \widetilde{\sigma_B^k} < \sigma < \widehat{\sigma}_N^k$, the bourgeoisie can only impose part of the investment on the landowners, namely $\overline{I_N^k} \geq 0$ and has to finance the rest of the payment $\widetilde{I_B^k}$. Instead, for $\sigma < \widetilde{\sigma_N^k} = \widetilde{\sigma_B^k}$ education is not provided by the elites.

In turn, Figure 2 represents the outcome for $H^k > 2$, a situation in which the payoffs from education for the bourgeoisie relative to the landowners are particularly high. In this case, the elite is willing to provide education if and only if $\sigma > \widehat{\sigma}_B^k$. The main difference with the preceding case is that for $\widehat{\sigma}_B^k < \sigma < \underline{\sigma}_N^k$, the bourgeoisie is willing to provide education even if it has the bear the full burden. In addition, in this area, the landowners become actually worse-off after the implementation of education.

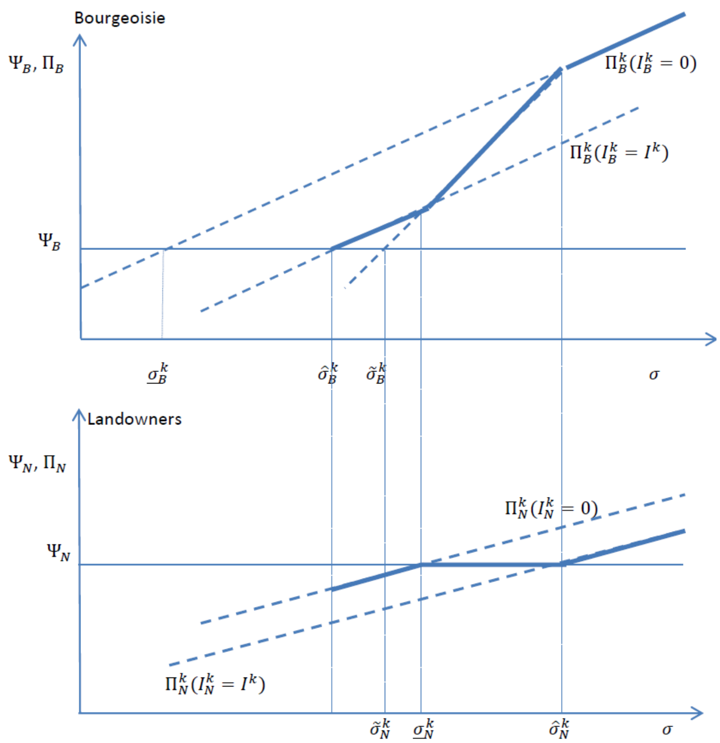


Fig. 2: Bourgeoisie dominant and $H^k > 2$

2.4.2 Landowners dominant

Figure 3 represents the case where the landowners are dominant and $H^k < 2$. In this case, the elite is willing to provide education if and only if $\sigma > \widetilde{\sigma}_N^k$. This provision is fully financed by the bourgeoisie if $\sigma > \widehat{\sigma}_B^k$ and partially financed by each group otherwise, i.e. the payments are $\widetilde{I_N^k}$ and $\overline{I_B^k}$ for respectively landowners and bourgeois.

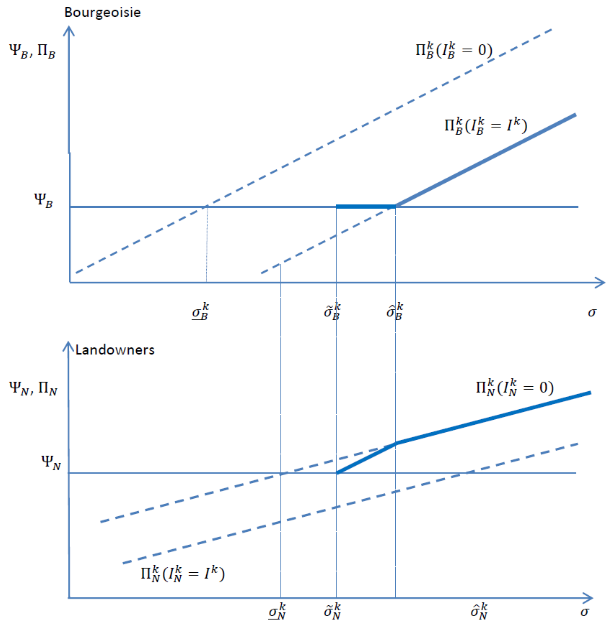


Fig. 3: Landowners dominant and $H^k < 2$

For $H^k > 2$, education is provided if and only if $\sigma > \underline{\sigma}_N^k$ and always fully funded by the bourgeoisie.

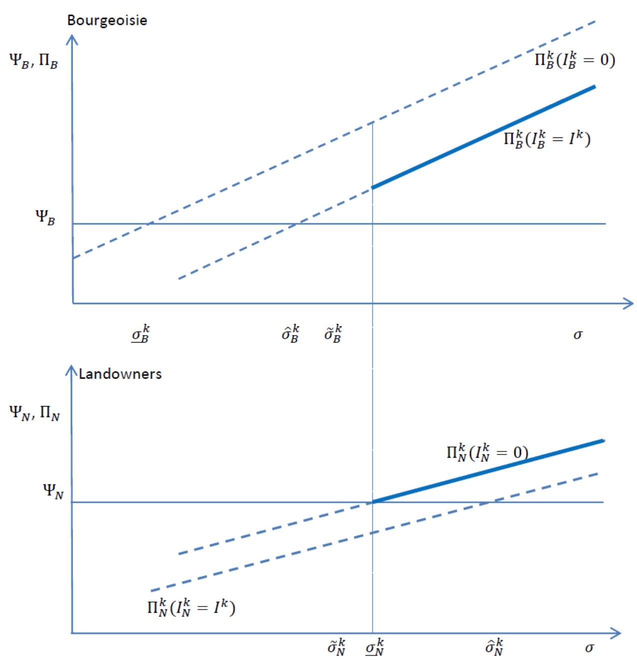


Fig. 4: Landowners dominant and $H^k > 2$

A simple look at the figures reveals that for $H^k < 2$ the elite agrees when to provide education (Figures 1 and 3). However, for $H^k > 2$ (Figures 2 and 4), the bourgeoisie is willing to fully finance education when the landowners do not even want education ($\widehat{\sigma}_B^k < \underline{\sigma}_N^k$), hence the bourgeoisie will provide education earlier than the landowners if the masses follow suit. We next study whether the masses want to follow suit and get schooled voluntarily.

2.5 School attendance by the masses

The masses of region i are willing to get educated whenever the payoffs from schooling are higher than the payoffs from no-schooling, i.e. $\Pi_{M_i} \geq \Psi_{M_i}$. Equalizing (2) and (3), the productivity threshold determining whether region-i masses attend schools under unified schooling is:

$$\sigma_{M_i}^{U} = \frac{1}{\delta} - \frac{2\left(\mu B - B_i + \delta B_{-i}\right)}{\delta\left(N + 2\mu B\right)} \qquad \text{for } i = 1, 2. \tag{15}$$

Similarly, from (2) and (6), the threshold for region-i schooling is:

$$\sigma_{M_i}^{R_i} = \frac{1}{\delta} - \frac{2B_i\left(\mu - 1\right)}{\delta\left(N + 2\mu B_i\right)} \qquad \text{for } i = 1, 2. \tag{16}$$

It is easy to show that, due to the increased match pool, the masses from each region are willing to get schooled earlier under unified than under regional schooling, i.e. $\sigma_{M_i}^{R_i} > \sigma_{M_i}^{U}$ for $i = 1, 2$.

At the same time, from both (15) and (16), it appears that the masses of i and $-i$ do not exhibit the same willingness to attend school even within a given educational system. Specifically, the masses of the region with a larger bourgeoisie have a higher productivity cutoff under unified schooling and instead a lower cutoff under regional schooling, i.e., it can be shown that $\sigma_{M_1}^{U} > \sigma_{M_2}^{U} \Leftrightarrow \sigma_{M_1}^{R_1} < \sigma_{M_2}^{R_2} \Leftrightarrow B_1 > B_2$. The underlying intuition is as follows: under unified schooling, the masses can get matched to the bourgeoisie of both regions, and hence the increase in the match pool is larger for the masses belonging to the region with a smaller bourgeoisie, which explains why they are willing to get schooled sooner. Instead, under regional schooling, the match pool is unchanged after education and thus the productivity gain stemming from schooling is larger for those masses which have already access to a larger bourgeoisie.

Since unified schooling requires the masses of both regions to be willing to get educated, the cutoff of the masses that are less willing to get schooled, namely the masses of region 1, $\sigma_{M_1}^{U}$ determines when unified schooling is feasible for the masses. In addition, given that $\sigma_{M_1}^{U} < \sigma_{M_1}^{R_1}$, regional education of the masses is never possible before unified schooling.

2.6 Equilibrium education

Lemma 2 shows that the incentives of the masses are irrelevant for the implementation of schooling:

Lemma 2 *Education always pays off for the masses when it does for the elite.*

Proof. See appendix B. ■

It is easy to understand why schools benefit the masses more than landowners even when the latter do not have to pay for setting up the schools: both groups require that their first period production loss due to schooling is offset by their gain due to higher productivity in the second period. While this productivity gain is the same for both groups when a member of the masses is matched to a landowner, a member of the masses might also be matched to a bourgeois granting him an even higher rise in productivity. In addition, under unified schooling, this advantage for a member of the masses is amplified by the possibility to be matched with a member of the bourgeoisie in both regions.

A parallel argument runs for the incentives of the bourgeois relative to those of the masses. Indeed, on the one hand, a match between a bourgeois and a mass member is the most productive match and the prospect for a mass member to be matched with some probability to a landowner puts initially its productivity threshold above that of the bourgeois. However, for such low productivity shocks, the landowners are not willing to finance schooling, and hence the bourgeoisie has to fully finance the setup of schools, which pushes their minimum productivity shock above the cutoff of the masses. Therefore, the incentives of the elites alone determine the implementation of schooling.

Proposition 1 *For $H^k < 2$ schooling is implemented for $\sigma > \widetilde{\sigma}_e^k$ independently of the identity of the dominant group. For $H^k > 2$ schooling is implemented earlier (specifically, for $\sigma > \widehat{\sigma}_B^k$) when the bourgeoisie is dominant than when landowners are dominant (implemented for $\sigma > \underline{\sigma_N^k} > \widehat{\sigma}_B^k$).*

Proof. Follows directly from the analysis in Section 2.4. B and Lemma 2. ■

We now discuss how the implementation of schooling varies with the underlying parameters.

When $H^k < 2$ the interests of both elite groups coincide. The cutoff for schooling to be implemented $\widetilde{\sigma}_e^k$ is decreasing in μ and B (or B_i respectively) and increasing in β.[10] A higher productivity advantage of the bourgeoisie μ makes schooling more profitable for the bourgeoisie reducing its threshold to implement schooling when dominant. But the threshold is also lower when landowners are dominant since they can make the bourgeoisie pay a larger amount of the schooling costs. Similarly, since the match productivity of the bourgeoisie increases more than that of landowners if the masses are schooled, an increase in the size of the bourgeoisie makes schooling more attractive. Finally, a higher share of the match productivity for the masses β raises the cutoff for schooling as the elites get a smaller piece of the cake in that case.[11]

For $H^k > 2$ the implementation of schooling depends on the identity of the dominant group when $\underline{\sigma_N^k} > \sigma > \widehat{\sigma}_B^k$. Landowners do not benefit from schooling and will not implement schools if dominant while the bourgeoisie benefits that much that it is willing to fully finance schools if it is in a position to do so. Again, the cutoff for a dominant elite to implement education $\widehat{\sigma}_B^k$ is decreasing in μ and and B (or B_i respectively) and increasing in β.[12]

[10] This can be shown by simple algebra.

[11] In this sense, in our model higher pre-industrial wages make early industrialization unlikely when the new technology is still fairly inefficient while elite groups who face lower pre-industrial wages might already implement industrialization.

[12] The threshold for dominant landowners $\underline{\sigma_N^k}$ is independent of these parameters: they never have to co-finance education (therefore the cutoff is independent of μ) and hence are willing to implement schooling when the second period productivity gain due to the technological advantage outweighs the cost of the loss of production in the first period. Since they get the same share of both the loss and the productivity gain, the cutoff is independent of β.

So far we have taken the potential educational system as given. Next, we turn to the choice of the education system by the dominant elite.

3 Unified vs. Region-i education

Under regional education, either region-1 or region-2 might become educated. The thresholds of all groups to get educated are weakly lower under region-1 than under region-2 education. Moreover, $H^{R_1} > H^{R_2}$. Therefore

Lemma 3 *The dominant elite always prefer region-1 schooling to region-2 schooling.*

Proof. See appendix C. ■

The intuition for this is simple: as the size of the landowners is the same in both regions, the productivity gains are larger when the masses with the larger bourgeoisie get educated. The bourgeoisie prefers this option as the return will be larger, and dominant landowners prefer it because it enables them to extract a larger payment from the bourgeois.

However, unified schooling is even better: the cost-savings of sending only one region to school do not outweigh the benefits from higher productivity in both regions and the increased match pool, as shown in the following lemma:

Lemma 4 *The dominant elite always prefers unified schooling to region$-i$ schooling.*

Proof. See appendix C. ■

It is easy to see that regional schooling in both regions is also dominated by unified schooling. Schooling both regions costs the same as unified schooling, but there is no regional mobility and hence the bourgeoisie loses out on the increased match pool across regions.

4 Secession

So far, we have assumed the existence of inter-regional transfers within elite groups leading to a perfect equalization of payoffs across regions within elite groups. In this section, we study whether the region-i dominant elite has actually incentives to avoid such redistribution by accompanying the implementation of schooling in region i by the political secession of this region. We assume that after region-i secession, no cross-border production can take place.

Since there are no interregional matches after secession, the cutoffs for the masses to be willing to go to school under region-i secession (denoted by S_i) are the same than under regional education, i.e. $\sigma_{M_1}^{S_1} = \sigma_{M_1}^{R_1} < \sigma_{M_2}^{S_2} = \sigma_{M_2}^{R_2}$. Instead, region-$i$ bourgeoisie's payoff from schooling with secession is

$$\Pi_{B_i}^{S_i} = -I_{B_i}^{S_i} + \delta(1-\beta)(1+\mu\sigma)\frac{M}{2} \tag{17}$$

i.e., the region-i bourgeoisie invests $I_{B_i}^{S_i}$ in the set-up of schools in its region and gets the proceeds from the future high-productivity matches with region-i masses. Similarly, the payoff from region-i secession for region-i landowners is:

$$\Pi_{N_i}^{S_i} = -I_{N_i}^{S_i} + \delta(1-\beta)(1+\sigma)\frac{M}{2}. \tag{18}$$

Equalizing (17) and (18) to (1), the productivity thresholds for the implementation of schooling with region-i secession for respectively region-i bourgeois and landowners are:

$$\sigma_{B_i}^{S_i} = \frac{2I_{B_i}^{S_i}}{\delta\mu(1-\beta)M} + \frac{1}{\mu\delta} \tag{19}$$

$$\sigma_{N_i}^{S_i} = \frac{2I_{N_i}^{S_i}}{\delta(1-\beta)M} + \frac{1}{\delta}. \tag{20}$$

Following the same steps as in section 2.3 and taking into account that educational costs are only paid by the regional elite, Table 1 in appendix A displays the cutoffs for free education, full payment and partial payment, together with the corresponding educational costs under S_i. It turns out that while educational costs differ, the cutoffs are the same as under regional schooling. Therefore Lemma 1 extends also to $k = S_i$ with $H^{S_i} = H^{R_i}$ and again we have two possible regimes depending on the profitability of schooling for the bourgeois relative to the landowners.

5 Secession versus unified schooling

We next study the choice between secession and unified schooling. For landowners, combining (5) and (18), we obtain that

$$\Pi_{N_i}^{S_i} \geq \Pi_N^U \Leftrightarrow I_{N_i}^{S_i} \leq I_N^U \tag{21}$$

i.e. landowners will simply go for the cheapest system in terms of investment, because they do not benefit from the extra cross-regional matches generated under unified schooling. This implies in particular that if they are to fully finance education under both systems, they will be indifferent between the two schooling systems as secession halves the number of mass members to be educated but also the number of landowners financing education, i.e. $I_{N_i}^{S_i} = \frac{M/2}{N/2} = I_N^U = \frac{M}{N}$.

Instead, secession compared to unified schooling restricts the number of matches for the bourgeois, which implies that secession will be preferred by the bourgeois only if it generates a sufficiently large reduction in costs. Note however that, as for landowners, the relevant cost is not the total expenditure in schooling, but the expenditure per member of the bourgeoisie: when going from unified schooling to region-i secession, the number of bourgeois financing education falls from B to B_i, which implies that the cost per bourgeois will not fall a lot unless B_i is very big. Mathematically, from (4) and (17), the condition under which secession is preferred is given by:

$$\Pi_{B_i}^{S_i} \geq \Pi_B^U \Leftrightarrow I_B^U - I_{B_i}^{S_i} \geq (1+\mu\sigma)(1-\beta)\delta\frac{M}{2}. \tag{22}$$

Clearly, as the costs of education are crucial in the secession decision and these costs partly depend on the identity of the dominant group, the choice between these two systems is likely to depend on the identity of the dominant group at the country and regional level. Subsections 5.1 and 5.2 study the equilibrium when respectively the bourgeoisie and landowners are dominant both at the country and regional level, while subsections 5.3 and 5.4 consider in turn the two cases in which the identity of the dominant group at the country and regional level is not the same.

5.1 Bourgeoisie always dominant

Proposition 2 shows that for a bourgeoisie dominant both at the regional and country level, cost saving from secession is never sufficient to offset the associated forgone productive matches.

Proposition 2 *A regionally and countrywide dominant bourgeoisie always prefers unified schooling to secession.*

Proof. See Appendix D.1 ■

In order to provide intuition for this result, Figure 5 compares for one of the three possible parameter configurations ($2 < H^{S_i} < H^U$) the payoffs from schooling under unified education (dashed line) and secession (continuous line) for the bourgeois and the landowners. For high enough productivity levels ($\sigma > \widehat{\sigma}_N^U$), the landowners are willing to pay for the entire cost of education under both systems, and thus the bourgeois choose unified schooling as the additional matches under unified schooling can be obtained at no extra cost. In turn, for $\underline{\sigma_N^k} < \sigma < \widehat{\sigma}_N^U$, the landowners are willing to pay the same amount of cost (per mass member) under both systems, and the rest needs to be paid for by the bourgeois. Then, as unified schooling is characterized by a larger set of matches for the bourgeois, the bourgeois' payoffs after the payment of these costs is higher under unified schooling. For lower productivity values ($\widehat{\sigma}_{B_i}^{S_i} < \sigma < \underline{\sigma_N^k}$), landowners are not anymore willing to contribute to education as this would make them worse-off, but bourgeois still implement schooling under both systems paying the full cost. Unified schooling is always preferred as the potential saving in terms of set-up cost stemming from secession –occurring only when the seceding region has a large bourgeoisie- is never sufficient to compensate for the loss of matches. Finally, for $\widehat{\sigma}_B^U < \sigma < \widehat{\sigma}_{B_i}^{S_i}$ schools are set up only under unified schooling, and so the bourgeoisie favours this system.[13]

[13] The two other parameter specifications are similar, except that there is no parameter area in which full payment by the bourgeoisie simultaneously arises as an equilibrium under both systems, and there is instead a situation in which partial payment by the bourgeoisie under unified schooling arises at the same time as no schooling under secession.

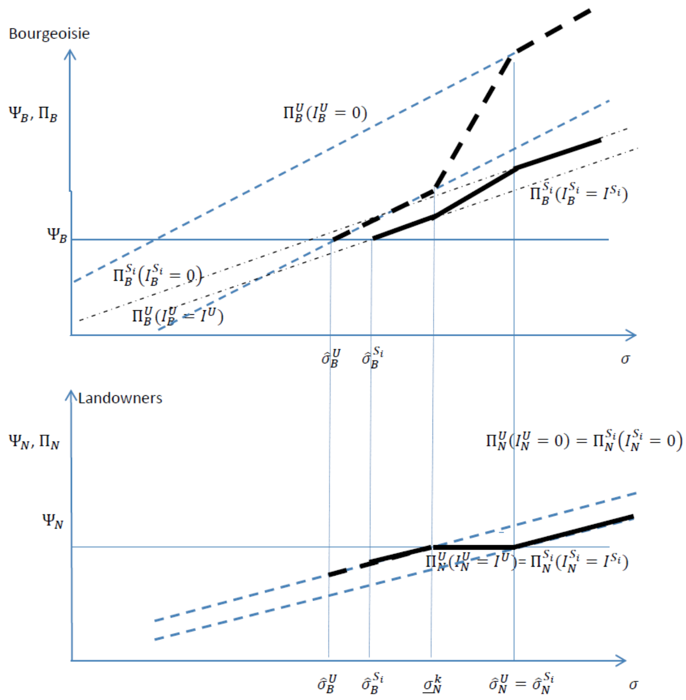


Fig. 5: Bourgeoisie dominant both at the country and regional level with $2 < H^{S_i} < H^U$

5.2 Landowners dominant always

As the payoff from schooling to landowners is the same under both systems, dominant landowners simply choose the system that allows them to transfer a larger share of the cost of schooling to the bourgeois. As the bourgeoisie benefits more from schooling under unified education, the bourgeoisie is willing to pay a larger share of the cost under this system, and landowners will always weakly prefer unified schooling to secession:

Proposition 3 *Regionally and countrywide dominant landowners always weakly prefer unified schooling to secession.*

Proof. See Appendix D.1. ■

Specifically, for high enough productivity levels, the bourgeoisie is willing to fully finance education under both systems, in which case landowners are indifferent between them. However, for intermediate productivity levels, full or even partial financing is only possible under unified schooling, and landowners will choose unified schooling for that reason.

5.3 Region-i-dominant but countrywide-dominated bourgeoisie

If the landowners are dominant at the country level but the bourgeoisie is dominant in region i, the bourgeoisie might want region i to separate. Specifically, the trade-off facing the bourgeoisie is as follows: on the one hand, if unified schooling can be implemented, secession leads to the loss of valuable match partners in region $-i$ (a loss that is increasing in $\mu\sigma$). On the other, the bourgeoisie can shift educational costs to the landowners under secession while it bears most of the costs under unified schooling as it is dominated by the landowners under that system. Hence if secession stands a chance against unified schooling, it has to be for relatively low productivity shocks – otherwise, the loss of potential partners would be too costly- but still high enough for education under secession to be profitable for the bourgeoisie. This intuition is confirmed in Proposition 4 which characterizes the equilibrium outcome since the masses are willing to get educated whenever the elite is willing to implement education.

Proposition 4 *A region-i dominant but country-level-dominated bourgeoisie chooses region-i secession if and only if (i) schooling is implemented under secession but not under unified schooling, or (ii) the productivity shock σ takes intermediate values and some further conditions are satisfied (see Appendix D.2.1 for the specific values and conditions).*

Proof. See Appendix D.2.1 ■

Part (i) in the proposition applies when schooling under secession is so beneficial for the regionally dominant bourgeoisie that it is profitable for them to implement it even when paying the full cost schooling. Instead, the landowners are made worse off by schooling and for this reason they choose not to implement schooling if they dominate under unified schooling.[14] However, once unified schooling becomes implementable, the bourgeoisie prefers being dominated under unified schooling over being dominant under secession due to the increased match pool under unified schooling.

As for Part (ii), Figure 6 illustrates one case where there exists a range of intermediate productivity shocks for which region-i bourgeoisie favours secession.[15] Indeed for low enough productivity levels, unified schooling is preferred either because schooling is simply not profitable under secession (for $\widetilde{\sigma}_B^U < \sigma < \widetilde{\sigma}_{B_i}^{S_i}$) or because landowners' willingness to pay under secession is limited (for $\widetilde{\sigma}_B^{S_i} < \sigma < \sigma_{aa2}$).[16] Symmetrically, for high enough productivity levels ($\sigma > \sigma_a$),[17] unified schooling dominates as the gain associated to having additional production partners is very high. Instead, for intermediate values ($\sigma_{aa2} < \sigma < \sigma_a$), region-$i$ bourgeoisie chooses to secede because it has to pay little for education under secession compared to unified schooling ($I_B^U = I^U$ while $I_B^{S_i}$ is either small or zero) and this actually outweighs the higher production under unified schooling.

[14] This happens for $2 < H^{S_i}$ when $\widehat{\sigma_{B_i}^{S_i}} < \sigma < \underline{\sigma_N}$.

[15] This case holds when $2 > H^{S_i} > H^U$, $B_i < \frac{N}{2\mu}$ and $\widehat{\sigma}_B^U < \widetilde{\sigma}_B^{S_i}$ are simultaneously verified.

[16] As shown in Appendix D.2.1 the cutoff σ_{aa2} is defined by $\Pi_B^U\left(I_B^U = \frac{M}{B}\right) = \Pi_{B_i}^S(I_{B_i}^{S_i} = \widetilde{I_{B_i}^{S_i}})$.

[17] As shown in Appendix D.2.1 the cutoff σ_a is defined by $\Pi_B^U\left(I_B^U = \frac{M}{B}\right) = \Pi_{B_i}^S(I_{B_i}^{S_i} = 0)$.

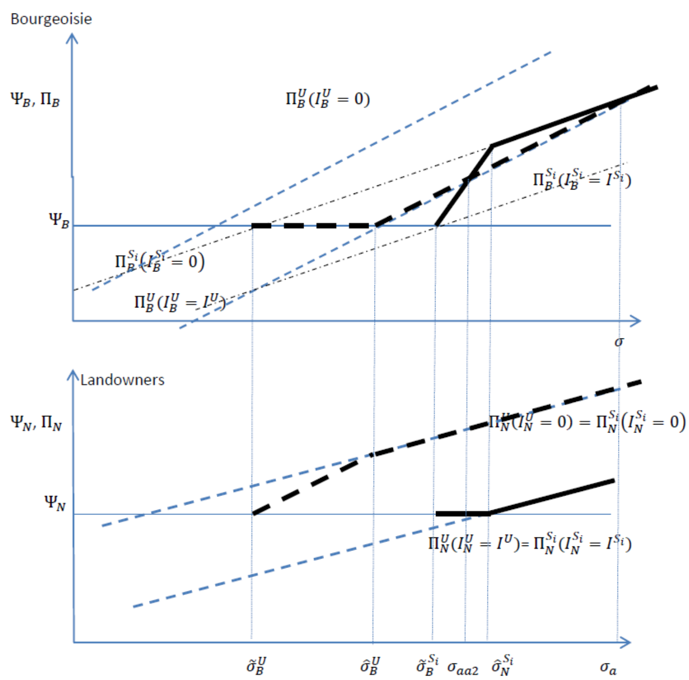


Fig. 6: Landowners dominant under U and bourgeoisie dominant under secession, $H^{S_i} < H^U < 2$

More generally, the specific range of intermediate productivity shocks for which secession is chosen depends on the size of region-i bourgeoisie. Indeed, while only half of the masses get educated under secession, the associated *per bourgeois* cost is smaller the larger the bourgeoisie of the seceding region. Other parameters which make the results case specific are the regional size of the bourgeoisie compared to the landowners –since the bourgeoisie shifts educational costs to the landowners- and the productivity advantage of the bourgeoisie.

While the bourgeoisie might prefer secession, the landowners never prefer to be dominated under secession to being dominant under unified education. Indeed, as discussed above, from the landowners' viewpoint, the only difference between the two systems are the educational costs and these are always higher under secession.

5.4 Region-i-dominant but countrywide-dominated bourgeoisie

Since the landowners do not benefit from regional mobility, they prefer secession whenever their educational costs under secession are lower than under unified education. This indeed happens if landowners are dominated at the country level but dominant in region i.

Proposition 5 *Region-i dominant but countrywide-dominated landowners always prefer region-i secession whenever education is implementable under secession. Hence only for productivity shocks for which unified education is implementable but education under secession is not, do we observe unified education.*

Proof. See Appendix D.2.2. ■

The landowners prefer secession because they are the dominant group under secession and therefore can shift (part of) the educational costs to the bourgeoisie and hence implement schooling paying less than they would under the unified system where they are the main bearers of the educational cost.[18]

We are now in a position to summarize our results. Secession can only be an equilibrium outcome if it implies a change in the dominant group. It will always result when education under secession is implementable and the landowners are dominant under secession but dominated at the country level. If it is the bourgeoisie that is dominant under secession but dominated at the country level, secession might occur only for intermediate productivity shocks. In the latter case it will also occur when schooling under secession is implementable and fully financed by the dominant bourgeoisie while unified schooling does not occur since the landowners are worse off under schooling.

We next study the welfare properties of the equilibria.

6 Welfare

The value of welfare in our model is obtained by adding up individual utility levels. In the absence of schooling, welfare is given by

$$W^{NS} = \frac{M}{2}\Psi_{M_i} + \frac{M}{2}\Psi_{M_{-i}} + B\Psi_B + N\Psi_N,$$

which using (1) and (2) simplifies to:

$$W^{NS} = \frac{M}{2}(1+\delta), \tag{23}$$

i.e. in every period elite members are matched to the masses of their region and produce one unit of output. How the production is split is a simple transfer from one group to the other and for this reason does not enter the expression.

Under unified schooling, welfare becomes:

$$W^U = \frac{M}{2}\Pi^U_{M_i} + \frac{M}{2}\Pi^U_{M_{-i}} + B\Pi^U_B + N\Pi^U_N,$$

which simplifies to

$$W^U = -M + \left((1+\sigma)\frac{N}{2} + (1+\mu\sigma)B\right)\delta M \tag{24}$$

using (3), (4), and (5). The direct cost of unified schooling is M, as all mass members get educated. As for the benefits, productivity is now higher ($1+\sigma$ and $1+\mu\sigma$ in the matches with respectively landowners and bourgeois) and the masses can now produce with the entire bourgeoisie. However, no production takes place in the first period as the masses are attending school during that period.

[18] The bourgeoisie never prefers being dominated under secession to being dominant under unified schooling as secession implies it loses valuable match partners and in addition in this case the bourgeoisie becomes the main bearer of educational costs under secession.

The social planner prefers unified schooling to no schooling if and only if $W^U > W^{NS}$ or equivalently when

$$\sigma > \sigma_W^U = \frac{2 - B(2\mu + \delta - 1)}{\delta(N + 2B\mu)} + \frac{1}{\delta} \tag{25}$$

This cutoff (25) decreases in δ, B and μ as schooling is more beneficial the more the future matters, the bigger the bourgeoisie, and the higher the productivity gains.

Next, we study the efficient educational level under regional schooling and secession. Clearly, as the planner maximizes utilitarian welfare and the payoffs of the agents are linear, distributional issues are irrelevant, and the implementation of schooling in one region and the secession of that region are indistinguishable from a welfare viewpoint. If region i is the only region that gets educated, welfare under secession or regional schooling is given by

$$W^{R_iS_i} = \frac{M}{2}\Pi_{M_i}^{R_i} + \frac{M}{2}\Psi_{M_{-i}} + B\Pi_B^{R_i} + N\Pi_N^{R_i},$$

which simplifies to

$$W^{R_iS_i} = -\frac{M}{2} + \frac{M}{2}(1+\delta)\left(B_{-i} + \frac{N}{2}\right) + \delta\frac{M}{2}\left(\frac{N}{2}(1+\sigma) + B_i(1+\mu\sigma)\right): \tag{26}$$

the direct cost of education is now given by $\frac{M}{2}$, while one unit of the good is produced in each period by each of the $\frac{M}{2}$ members of the masses in region $-i$ with the regional elite $B_{-i} + \frac{N}{2}$ and production in region i is confined to the second period but with a higher productivity ($1 + \sigma$ with the landowners and $1 + \mu\sigma$ with the bourgeois).

Subtracting (23) from (26), the planner prefers education in one region to no education if and only if:

$$\sigma > \sigma_W^{R_iS_i} = \frac{1 + B_i + \frac{N}{2}}{\delta\left(\mu B_i + \frac{N}{2}\right)}. \tag{27}$$

6.1 Efficiency of equilibrium education

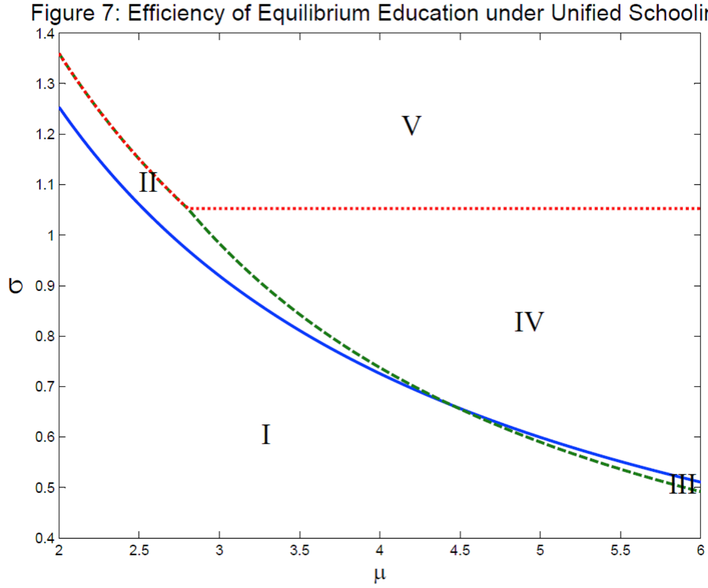
Proposition 6 checks whether the implementation of unified schooling in the decentralized equilibrium is socially efficient:

Proposition 6 *The equilibrium education level under unified schooling can be socially inefficient. Undereducation arises under a wide set of parameters, while overeducation can arise only if the bourgeoisie is dominant, the payoffs from schooling for the bourgeoisie are high relative to the payoffs of the landowners and the productivity takes intermediate values as specified in Appendix E.2.*

Proof. See Appendix E.2 ■

The main intuition behind the inefficient provision of unified schooling is simply that the politically dominant elite does not internalize the benefits from schooling for the other elite and for the masses. More specifically, the solid line in Figure 7 represents the socially efficient productivity threshold (σ_W^U) for the provision of education for different values of the productivity advantage in matches involving the bourgeoisie (μ) while the dashed (resp. dotted) line represents the equilibrium threshold under unified schooling when the bourgeoisie (resp. landowners) are the dominant group.[19]

Figure 7: Efficiency of Equilibrium Education under Unified Schooling



Equilibrium education under the unified system is efficient no matter the identity of the dominant group in area I (no education) and area V (education), while in area IV the efficient choice is made only if the bourgeoisie is the dominant group. In area II, education is efficient, but is never implemented because the bourgeois do not take into account the gains education generates for landowners. Instead, in area III there is actually overeducation under unified schooling if the bourgeoisie dominates. Specifically, the bourgeois fully finance education and make landowners worse-off than under no education as the productivity gain for landowners is too low to cover for the loss of first period production while the masses take education. As this loss is not internalized by the bourgeois, the equilibrium is characterized by excess education.

The same type of argument holds regarding the efficiency of decentralized education under region-i schooling or secession, but overeducation arises only for values of the bourgeoisie's productivity advantage (μ) higher than under unified schooling because the bourgeoisie does not gain production partners under these systems. Figure A1 in the appendix provides an example of these different cases.

6.2 First best

Assume first that the planner is able to enforce the welfare maximizing schooling level under each system.[20] In that case, the planner prefers unified schooling over regional schooling/secession if

[19] We have assumed that $\beta = 0.1$, $\delta = 0.95$, $B_1 = 0.25$, $B = 0.4$, and $M = 2$. The threshold for a dominant bourgeoisie (dashed line) is given by $\widetilde{\sigma}_B^U$ for $H^U < 2$ and by $\widehat{\sigma}_B^U$ for $H^U > 2$ and the threshold for landowners (dotted line) by $\widetilde{\sigma}_B^U$ for $H^U < 2$ and by $\underline{\sigma}_U^N$ for $H^U > 2$ (see Figures 1 to 4).

[20] As it is easy to show that $\sigma_{M_2}^U < \sigma_{M_1}^U < \sigma_W^U$ and $\sigma_{M_i}^{R_i} = \sigma_{M_i}^{S_i} < \sigma_W^{R_i S_i}$ for $i = 1, 2$, the masses are always willing to attend school whenever the central planner wants this to be the case.

and only if $W^U > W^{R_i S_i}$ $\forall i = 1, 2$. As it is easy to show that $W^{R_1 S_1} > W^{R_2 S_2}$,[21] the condition becomes $W^U > W^{R_1 S_1}$, which from (24) and (26) is equivalent to:

$$\sigma \left(\frac{N}{2} + \mu B_2 \right) \frac{M}{2} \delta + (1 + \mu\sigma) B \frac{M}{2} \delta > \frac{M}{2} + \frac{M}{2} \left(B_2 + \frac{N}{2} \right). \tag{28}$$

The social planner implements unified schooling whenever the additional returns from education in this system compared to regional schooling/secession (LHS of (28)) outweigh the additional costs of unified schooling (RHS). Specifically, the additional returns come from the higher productivity $\sigma \left(\frac{N}{2} + \mu B_2 \right) \frac{M}{2} \delta$ of the now-educated region-2 masses in their matches with region-2 elites and also from the additional matches the unified system enables, namely the inter-regional production between the masses of region i and the bourgeois of region $-i$ (for $i = 1, 2$), i.e. $(1 + \mu\sigma) B \frac{M}{2} \delta$. In turn, unified schooling is more costly because it educates $\frac{M}{2}$ additional mass members, and also because region-2 masses do not produce anymore in the first period $\frac{M}{2} \left(B_2 + \frac{N}{2} \right)$ as they are attending school.

Clearly, for sufficiently high values of μ and/or σ in (28), the planner will choose unified schooling. The question then remains whether the planner may choose region-i schooling when the returns from education are sufficiently low. Proposition 7 shows that this will never be the case, because welfare under region-i schooling is higher than welfare under the unified system only when implementing schooling is actually inefficient. Intuitively, in that case, given that education is inefficient, implementing it a lower scale (region-i schooling) is better than implementing it at a higher scale (unified schooling). However, as the planner has always the possibility of not implementing school at all, implementing region-i schooling is never a first best outcome:

Proposition 7 *Under centralization, unified schooling yields higher welfare than regional schooling and secession.*

Proof. See appendix E.1 ■

6.3 Second best

In general, the central planner will not be able to control the school set-up investments made by landowners and bourgeois. If the elites are willing to finance unified schooling, this system will still be chosen by the central planner. However, if the elites are not willing to implement that system and they are instead willing to implement regional schooling or secession, the central planner may prefer this option if for the corresponding parameter area education in one region is socially better than no education.

[21] Indeed, the planner prefers the implementation of schooling in the region with the larger bourgeoisie (region 1) because (i) the cost of schooling is the same in both regions given that the size of the masses is identical, and (ii) education increases more the productivity of matches with the bourgeoisie, and thus the productivity gains from education are higher for those masses with access to a larger bourgeoisie. Mathematically, $W^{R_1 S_1} > W^{R_2 S_2}$ holds if $\sigma > \frac{1}{\delta\mu} = \underline{\sigma_B^{R_i}}$, which is always satisfied.

From Propositions 2 and 3, we know that if the identity of the dominant group is the same countrywide and regionally, the dominant group always chooses unified schooling, which implies that regional schooling/secession can never be a second best in that case.

Consider next the case where the bourgeoisie is dominant at the country level but landowners are dominant at the regional level. From Proposition 5, landowners always prefer secession when both unified education and secession are implementable and in addition unified schooling is always implementable when education under secession is. As the central planner always prefers unified schooling when implementable, this means that any landowner-driven secession lowers welfare.

Finally, consider a regionally-dominant but countrywide-dominated bourgeoisie. In this case, we know from Proposition 4 that for $H^{S_i} > 2$ and $\widehat{\sigma}_{B_i}^{S_i} < \sigma < \underline{\sigma_N}$ schooling is implemented under secession but not under the unified system. This is not however a sufficient condition for secession to be a second best, because secession has still to be efficient, i.e. socially better than no-schooling. This condition does not automatically hold and requires the cut-off for efficient education in one region ($\sigma_W^{R_i S_i}$) to be lower than the minimum productivity level rendering schooling profitable for the landowners, i.e. $\underline{\sigma_N} = \frac{1}{\delta}$. This result is summarized in the following proposition:[22]

Proposition 8 *If a regionally-dominant bourgeoisie wishes to implement secession when the countrywide-dominant landowners do not want to implement unified schooling ($H^{S_i} > 2$ and $\widehat{\sigma}_{B_i}^{S_i} < \sigma < \underline{\sigma_N}$), secession is a second best if an only if* $\max[\sigma_W^{R_i S_i}, \widehat{\sigma}_{B_i}^{S_i}] < \sigma$.

Proof. See Appendix E.4 ■

7 Robustness

The above results are derived assuming one cross-regional dimension of heterogeneity, namely the size of the bourgeoisie. In this section we briefly discuss other forms of heterogeneity. As before we assume that the regions are identical except in one dimension. Specifically, we consider in turn cross-regional heterogeneity in the number of mass members, the number of landowners, or the date of the industrialization opportunity.

If the heterogeneity stems from the size of the masses, the minimum productivity shock necessary for the masses to be willing to get region$-i$ schooling or unified schooling becomes identical across regions. As before, the masses are willing to attend unified schools whenever they are willing to attend regional schools. Under region$-i$ education the elite of the region with the larger masses benefits more from education, but educating this region is also more costly since more individuals have to be schooled. Unified schooling then leads to less than double education costs and big benefits due to the mobility of the masses. This leads to nation building for sufficiently high productivity shocks and makes secession less likely than in our benchmark setting.

[22] In Appendix E.5 we provide one numerical example for which the secession of a regionally-dominated bourgeoisie is a second best, and one in which this outcome is worse than no-schooling.

It is easy to see that our results are robust to the case when it is the size of the landowners that differs across regions. The main difference to the benchmark model is that now the cutoffs for education (might) depend on the size of the regional landowners and no longer on the size of the regional bourgeoisie. Now the masses in the region with the smaller group of landowners are willing to get educated earlier in each educational system since they face a smaller probability to meet a landowner and are therefore are more likely to be matched with a bourgeois and enjoy the additional productivity increase of this type of match. Nevertheless, the masses are still willing to get education in the unified system before they are willing to get educated regionally or under secession. Moreover, as before, the masses are still willing to get educated whenever the elite is willing to implement education. Hence, the education incentives of the elite alone determine when schooling is implemented. They face the same type of trade-offs than before, and hence the results are qualitatively the same.

Finally, consider a case where only region i experiences a large industrialization shock, and as a result only the agents in that region can initially access the high technology. If attaining a high productivity in cross-regional matches requires only one of the agents –bourgeois or mass members- to have access to that technology, our results on the equilibrium and optimality of unified schooling still carry on. This is also the case if unified schooling creates a common identity that allows mass members to migrate across regions, following Gellner (1983). Instead, if migration is not possible and high productivity in cross-regional matches requires both the bourgeoisie to access the new technology and the masses to be trained for it, asymmetric industrialization shocks across regions hinder nation building and originate an additional channel leading to secession at equilibrium.[23] Indeed, in that case, regional schooling in the high-productivity region is more attractive for all groups than regional schooling in the low-productivity region. Moreover, the bigger the relative difference across regions, the more likely it is that regional education dominates unified schooling. In addition, since transfers to the less efficient region can be avoided by implementing secession, a very unequal speed of industrialization makes nation building across both regions impossible and is likely to lead to secession.

8 Case Study: Spain versus France

In this section, we relate our model to the cases of 19th century France and Spain. At the beginning of the 19th century, both countries were similar at least along some characteristics relevant to our model.

Specifically, both countries had a very similar per capita GDP at the beginning of the century[24] and were characterized by a heterogeneous language composition. Indeed, in 1794 only about 40% of the French population were native French speakers[25] (Calvet 2002, p. 218), while an important proportion of the Spanish population had a language other than Spanish/Castilian

[23] In the model this could be represented by different σ_i or by different μ_i.

[24] According to Tortella (1994, p. 2) Spain's per capita GDP was 2% higher than France's in 1800, and 7% lower in 1820. According to Maddison (2003, pp. 58-67), France' s GDP was higher by 11% in 1820.

[25] Among the other language groups, the largest was Occitan and next came Breton and Alsacian. Additionally, small minorities were speaking Franco-provençal, Basque, Catalan, Corsican, or Flemish.

(i.e. Catalan, Galician, Basque, or Bable) as their mothertongue in 1787.[26] Another common characteristic is that the first industries were geographically concentrated. In France, the first industries were mostly concentrated in the North-East (Crayen and Baten, 2010), and in the case of Spain they were mostly concentrated in Catalonia and in the Basque Country (Tortella, 2000).

Despite these common features, France and Spain ended up having very different outcomes in terms of industrialization, with France's per capita GDP becoming 1.7 times that of Spain in 1930.[27] The outcomes were also very different in terms of nation-building. In the historical literature, France is often used as a benchmark of successful nation-building (see e.g. Kroneberg and Wimmer, 2012) while Spain is seen as an example of an unaccomplished nation-building process accompanied with the emergence of peripheral nationalisms (see e.g. Linz, 1974, 1975; Keating, 1993). When elections were held, peripheral nationalist parties were systematically represented in the Spanish Parliament since the end of the 19th century. Instead, the success (or even the existence) of regionalist/nationalist parties in Alsace, Brittany, Corsica, or the French parts of the Basque Country or Catalonia has been extremely limited. For instance, in the June 1931 Spanish legislative elections, the Catalan nationalist parties obtained almost three fourths of the Catalan constituencies, and their Galician regionalist and Basque nationalist counterparts respectively 40 per cent and one third of the Galician and Basque constituencies (see Tusell, 1982).[28] In contrast, in the first round of the April 1928 French legislative elections, regionalist candidates were only present in Alsace and obtained 4 seats with 15.9% of the votes: overall, the French Parliament consisted of 4 regionalist deputies out of 612 (see Lachapelle, 1928).

In terms of our model, we can consider that the two regions characterizing France are the industrializing North-East and the agricultural South-West, as defined for instance by the "St-Malo-Geneva line" identified by some Historians (see e.g. Weber, 1976). In addition, it is safe to assume that the bourgeoisie was the dominant elite both in the North-East and at the French level as a whole. Indeed, Price (2004) argues that while the landowners retained an important amount of power at least until 1870, "'New' wealth was represented by a *grande*

[26] There are no available data on the language composition of Spain at the end of the 18th century. However, one can do a back of the envelope computation to get an upper bound for the proportion of non-Spanish speakers. According to Linz (1975), historically Spanish has also been spoken by part of the population in those regions where Catalan, Galician, or Basque were also spoken. Instead, these three languages were geographically concentrated in certain provinces (Barcelona, Tarragona, Lleida, Girona, Valencia, Castellon, Alicante, and the Balearic Islands, for Catalan; A Coruna, Lugo, Ourense, and Pontevedra, for Galician; and Gipuzkoa, Bizkaia, Araba, and Navarre for Basque). Using data on the population of provinces in the 1787 Census (INE, 1991), an upper bound for the proportion of Catalan, Galician, Basque, and Bable speakers is respectively 18%, 13%, 5%, and 3%, and thus a lower bound for the proportion of Spanish speakers is 61%.

[27] Measured in 1970 U.S. dollars adjusted for purchasing power parity, the GDP per capita of France in 1930 was 1,337 and that of Spain 798 (Tortella, 1994, p.2). Similar results are found in Maddison (2003, pp. 62 & 68): measured in 1990 international Geary-Khamis dollars, France's per capita GDP was 4,532 and Spain's 2,620.

[28] In Catalonia, out of 53 seats, *Esquerra Republicana de Catalunya* obtained 31 seats, the *Lliga Regionalista* 3, the *Unió Socialista de Catalunya* 2, the *Partit Català Republicà* 1, and the *Esquerra Catalana Radical-Socialista* 2. In Galicia, out of 47 seats, the *Federación Republicana Gallega* obtained 14, the *Galleguistas* 2, and the *Regionalistas* 1. In the Basque Country (excluding Navarre), out of 24 seats, the *Partido Nacionalista Vasco* obtained 8 seats.

bourgeoisie, which had, since 1830, achieved dominance not only in commerce, industry, and the professions but also in government" (Price, 2004, p. 37). As the bourgeoisie is dominant both at the regional and at the country level, from Proposition 2 we expect the bourgeoisie to choose unified schooling, which in turns results in the creation of a common French identity.

The implementation of schooling throughout the country and the creation of a strong common French identity were actually observed. According to Nuhoğlu Soysal and Strang (1989), while France introduced compulsory education only in 1882, the primary enrollment ratio was already 75 percent in 1870 (p. 278), the highest amongst developed countries. However, there were big cross-regional differences in school attendance (Weber, 1976). In the 1880s, schooling became free, French was made the only language of instruction (Chervel, 1992) and "village teachers, trained to greater competence and new self-respect, became the licensed representatives of the Republic" (Weber, 1976, p. 318). Parallel to this, parents started to perceive that numeracy and literacy were actually useful (as e.g. they were required to get jobs both in the public and the private sector), attendance increased, and differences in attendance across regions started to decline (Weber, 1976). At the same time, in Weber's (1976, p. 332) words, "the greater function of the modern school (is) to teach not so much useful skills as a new patriotism beyond the limits naturally acknowledged by its charges. The revolutionaries of 1789 had replaced old terms like schoolmaster, regent, and rector, with *instituteur*, because the teacher was intended to *institute* the nation". The successful implementation of the schooling system throughout the country constituted a "wide-ranging process of standardization that helped create and reinforce French unity, while contributing to the disintegration of rival allegiances" (Weber, p. 338).[29]

In the case of Spain, the two regions can be identified as the industrializing periphery (Catalonia and the Basque Country) and as the agricultural "centre" which comprises the rest of the country. According to historians, the bourgeoisie was dominant in the periphery, while the landowning elite from the centre dominated Spanish politics (see e.g. Linz, 1974; Solé Tura, 1989; or Harrison, 1990).[30] In our model, the case where the landowners are dominant at the country level and the bourgeoisie is dominant at the regional level is studied in Proposition 4. In this Proposition, secession arises as an equilibrium outcome when the industrialization shock is weak and is more likely to arise when the overall size of the bourgeoisie is small.[31] While we cannot directly observe the size of the industrialization shock, Keating (1993) and Balfour (1995) argue that the Catalan textile industry was uncompetitive by European standards, and required for this reason a protected market for its goods (Spain) and a protected source of raw materials (cotton from Cuba). Assuming that the overall size of the bourgeoisie is small at the

[29] Weber (1976, p. 336) also argues that "Teachers taught or were expected to teach 'not just for the love of art or science...but for the love of France – a France whose creed had to be inculcated in all unbelievers. A Catholic God, particularist and only identified with the fatherland by revisionists after the turn of the centurym was replaced by a secular God: the fatherland and its living symbols, the army and the flag. Catechism was replaced by civics lessons. Biblical history, proscribed in secular schools, was replaced by the sainted history of France".

[30] Harrison (1976, p. 902) argues for instance that "the agrarian and financial interests of central and southern Spain [who] made up the political oligarchy".

[31] Observe that secession is possible in more cases when $2 > H^U = (1-\beta)(2\mu - 1 + \delta)B$ which is easier to satify the smaller B.

Spanish level seems reasonable, given the very limited development of industries in the rest of the country, as underlined by the literature on the failure of the industrial revolution in Spain during the 19th century (see Nadal, 1973).

As predicted by the model, the development of the education system was weak in Spain and peripheral nationalisms developed, although secession was not observed. Indeed, while education became compulsory already in 1838, the primary enrollment ratio in 1870 was only of 42 percent (Nuhoğlu Soysal and Strang, 1989, p. 278) and "[c]entral government funding for primary education remained minimal: between 1850 and 1875 education never accounted for more than 1.13 percent of the budget and by the 1870s it had fallen to 0.55 percent" (Shubert, 1990, p. 182). The illiteracy rate was 71 percent in 1870 and still 50 percent in 1910, against respectively 32 and 13 percent in France in the same dates (Tortella, 2000, p. 13). In addition, Shubert (1990, p. 183) argues that "the war against non-official languages in Spain was much less successful in Spain than in France (...) One reason for this was that the Spanish state was much less effective in creating the basic agent of linguistic uniformity: the schools".

As for the development of Catalan nationalism, Linz (1974) argues that a regionalist movement started in Catalonia in the mid 19th century and turned into a nationalist movement at the end of the century. In addition to the cultural and literary revival of the Catalan language, "it was the defense of the interests of the national bourgeoisie that activated manufacturers to create interest groups, organize meetings, write petitions, and contribute decisively to the founding of the *Lliga de Catalunya* in 1887" (p. 62) one of the first Catalanist parties. However, the "minority character of the industrial bourgeoisie of Catalonia, and later the Basque country, in the total Spanish social structure, and the impossibility for it to gain power at the center within the oligarchic liberal democracy of the Restoration [1870-1931], turned it away from the struggle for power in the Spanish state. Instead it aimed to secure power at the local and regional level and to build up support on the basis of cultural nationalism to bargain more effectively with the central government on economic issues particularly protectionism" (Linz, 1975, pp. 384-386). Two examples of conflicts between the Catalan bourgeoisie and the centre's landowning elite are the fight over tariffs after Cuba's independence in 1898 -with the Catalan bourgeoisie defending the elimination of tariffs on foreign grain and the imposition of tariffs on foreign textiles (see Harrison, 1990, or Díez Medrano, 1994)- and over the taxation of industrial profits during World War I (see e.g. Carr, 1980 or Enrlich, 1998). Moreover, the programme of the Catalan employers' group *Fomento del Trabajo Nacional* set up following Cuba's independence stressed the implementation of technical education as one of four main demands (Harrison, 1974) which indicates that they valued schooling. Finally, referring to the end of the 19th century and beginning of the 20th, Balcells (2013) argues that "Catalan nationalism exists and is salient in Spain because people were never massively educated under a strong a well-organized Spanish state" (p. 468).

9 Conclusion

This paper presents a Gellnerian model of industrialization and nation building emphasizing the key role of elites in shaping that process. As in Gellner (1964, 1983), the central link between industrialization and nation building goes through the double role of schooling as productivity enhancer and generator of a common identity. In addition, as in more recent contributions to the nation building literature (see in particular Breuilly, 1993; Roeder, 2007; Kroneberg and Wimmer, 2012), the observed outcome in terms of industrialization and nation building crucially depends on the nature of the interaction between elite groups with different (and sometimes diverging) interests.

In our two-region model, nation building through a unified educational system which brings a common identity to both regions is assumed to be superior to regional organizations of education because it expands output by enabling inter-regional production. However, this first best outcome is not always reached at equilibrium if the regionally-dominant elite is dominated at the country level. The intuition for this result is as follows: an elite which is dominant at both geographical levels can appropriate a large share of the cake at both levels, and thus goes for the implementation of institutions at the level where the cake is the largest, i.e. the country level. Instead, a regionally-dominant but countrywide-dominated elite may prefer a large share of the small (regional) cake rather than a small share of the large cake stemming from nation-building. If the bourgeoisie is the regionally-dominant group, regional schooling through secession can actually be a second best when the milder interest for education of the countrywide-dominant landowners lead them not to embark in nation building.

As a historical check, we study the ability of the model to explain the divergent evolution of France and Spain in terms of nation-building and industrialization, despite their similar levels of income and linguistic diversity at the beginning of the 19th century. In France the bourgeoisie was dominant both at the country level and in the industrializing North-East; therefore according to our model we should expect nation-building. Instead the model predicts the absence of nation-building in Spain given that the peripheral (Catalan and Basque) bourgeoisie was only dominant at the regional level and the industrialization shock was not very large. Indeed, France experienced the systematic implementation of free primary education across the country since the 1880s, while Spain was still experiencing a 50% illiteracy rate in 1910. This very different investments in education ran parallel to the much higher growth and stronger national identity in France than in Spain.

While our model finds empirical backing in the comparative analysis of Spain versus France, our paper does not perform an econometric test of Gellner's theory. A important first step in this direction is Wimmer and Fenstein (2012) which uses a quantitative event history analysis, and finds a positive –but statistically insignificant- correlation between the length of railway tracks – thought of as a measure of nation-building through industrialization- and the birth of nation-states. A further step in this direction which is beyond the scope of this paper and requires an important data collection effort would ideally link educational investments during industrialization and the presence of national identities -whether or not taking the form of an independent state.

A Cutoffs and educational costs for the elite

Let e (resp. $-e$) denote the dominant (resp. dominated) group and E (resp. $-E$) its size. Then, educational costs are split as follows:

- For very high productivity shocks, $\sigma > \max\left[\underline{\sigma}_e^k, \widehat{\sigma}_{-e}^k\right]$, $I_e^k = 0$, and schooling is entirely financed by the dominated group. Under unified schooling, each member of the dominated group pays $\frac{M}{-E}$ since the masses of both regions get educated. Under region-i schooling, the cost reduces to $\frac{M}{2(-E)}$.

- If $\max\left[\underline{\sigma}_e^k, \widehat{\sigma}_{-e}^k\right] = \widehat{\sigma}_{-e}^k$,

  - then for $\max\left[\widetilde{\sigma}_e^k, \underline{\sigma}_{-e}^k\right] < \sigma < \widehat{\sigma}_{-e}^k$, the dominant group has to cofinance education paying $\widetilde{I_e^k}$ while the dominated group pays $\overline{I_{-e}^k}$. The value of $\widetilde{I_e^k}$ for the different systems is $\widetilde{I_e^U} = \frac{M - \overline{I_{-e}^U}(-E)}{E}$ and $\widetilde{I_e^{R_i}} = \frac{\frac{M}{2} - \overline{I_{-e}^{R_i}}(-E)}{E}$.
  - if $\max\left[\widetilde{\sigma}_e^k, \underline{\sigma}_{-e}^k\right] = \underline{\sigma}_{-e}^k$ and $\max\left[\underline{\sigma}_{-e}^k, \widehat{\sigma}_e^k\right] = \underline{\sigma}_{-e}^k$, then for $\widehat{\sigma}_e^k < \sigma < \underline{\sigma}_{-e}^k$, the dominant group wants education, but the dominated group is made worse off with education, so the dominant group fully pays the educational costs, namely $\frac{M}{E}$ under unified schooling and $\frac{M}{2E}$ under region-i schooling.

- In all other cases, the dominant group has no interest in implementing schooling.

Table 1 reports the values of the productivity thresholds and associated payments under unified schooling, region-i education and region-i secession.[32] In order to calculate the thresholds under secession, we take into account that in that case educational costs are paid only by the elites of the seceding region.

	Unified education	Region-i education	Region-i secession
$\underline{\sigma}_N^k$	$\frac{1}{\delta}$		
$\underline{\sigma}_B^k$	$\frac{1-\delta}{2\mu\delta}$	$\frac{1}{\delta\mu}$	
$\widehat{\sigma}_N^k$	$\frac{2}{(1-\beta)\delta N} + \frac{1}{\delta}$		
$\widehat{\sigma}_B^k$	$\frac{2-(1-\beta)B(2\mu-1+\delta)}{2\mu(1-\beta)\delta B} + \frac{1}{\delta}$	$\frac{1-(1-\beta)B_i(\mu-1)}{\mu(1-\beta)\delta B_i} + \frac{1}{\delta}$	
$\widetilde{\sigma}_e^k$	$\frac{2-(1-\beta)B(2\mu-1+\delta)}{(1-\beta)\delta(N+2B\mu)} + \frac{1}{\delta}$	$\frac{2-(1-\beta)2B_i(\mu-1)}{\delta(1-\beta)(N+2\mu B_i)} + \frac{1}{\delta}$	
$\overline{I_N^k}$	$\frac{(\delta\sigma-1)(1-\beta)M}{2}$	$\frac{(\sigma\delta-1)(1-\beta)M}{4}$	$\frac{(1-\beta)(\delta\sigma-1)M}{2}$
$\overline{I_B^k}$	$\frac{(1-\beta)[2\mu\delta\sigma-1+\delta]M}{2}$	$\frac{(\mu\sigma\delta-1)(1-\beta)B_iM}{2B}$	$\frac{(1-\beta)(\delta\mu\sigma-1)M}{2}$
$\widetilde{I_N^k}$	$\frac{2-B(1-\beta)(2\mu\delta\sigma-1+\delta)}{2N}M$	$\frac{1-(\mu\sigma\delta-1)(1-\beta)B_i}{N}\frac{M}{2}$	$\frac{1-B_i(1-\beta)(\delta\mu\sigma-1)}{N}M$
$\widetilde{I_B^k}$	$\frac{2-N(1-\beta)(\delta\sigma-1)}{2B}M$	$\frac{2-(\sigma\delta-1)(1-\beta)N}{B}\frac{M}{4}$	$\frac{2-N(1-\beta)(\delta\sigma-1)}{4B_i}M.$

Table 1: Productivity thresholds

[32] For the time being we ignore nonnegativity constraints on $\overline{I_e^k}$ when calculating $\widetilde{I_e^k}$ and $\widetilde{\sigma_e^k}$. This approach has the advantage that $\widetilde{\sigma_e^k} = \widetilde{\sigma_B^k} = \widetilde{\sigma_N^k}$ for all schooling systems k, but, as it might lead to unnatural rankings of the cutoffs, in particular to $\widehat{\sigma_B^k} < \widetilde{\sigma_B^k} = \widetilde{\sigma_N^k} < \underline{\sigma}_N^k$ when $H^k > 2$. This is of no importance, since $\widetilde{\sigma_e^k}$ is irrelevant in these cases and we therefore write this case as $\widehat{\sigma_B^k} < \underline{\sigma}_N^k$ only as in part 2 of Lemma 1.

B School attendance by the masses

Proof of Lemma 2 The minimum productivity shocks for the masses to be willing to be schooled are $\sigma_{M_1}^{U}$ (given by (15)) for unified schooling and $\sigma_{M_i}^{R_i}$ (given by (16)) for regional schooling. The relevant cutoffs for the elite are $\widetilde{\sigma}_B^k = \widetilde{\sigma}_N^k$ when $H^k < 2$, and $\underline{\sigma_N^k}$ for the landowners and $\widehat{\sigma}_B^k$ for the bourgeoisie when $H^k > 2$. We now show that in all cases these cutoffs are bigger than the cutoff of the masses to be willing to get schooled. In particular: (i) $\sigma_{M_1}^{U} < \widetilde{\sigma}_e^U$ always holds given that $\sigma_{M_1}^{U} < \widetilde{\sigma}_e^U \Leftrightarrow (1+\delta)(B_1 - B_2)(1-\beta) < 2$, and the second inequality always holds. (ii) $\sigma_{M_1}^{U} < \widehat{\sigma}_B^U$ always holds given that $\sigma_{M_1}^{U} < \widehat{\sigma}_B^U \Leftrightarrow 0 < \frac{2N}{(1-\beta)B} + 2\mu\left[\frac{2}{1-\beta} - 1 + 2(1+\delta)B_2\right] + (1-\delta)N$, and the RHS of the second inequality is always positive. (iii) $\sigma_{M_1}^{U} < \underline{\sigma_N^U}$ holds (trivial). (iv) $\sigma_{M_i}^{R_i} < \underline{\sigma_N^{R_i}}$ (trivial); (v) $\sigma_{M_i}^{R_i} < \widetilde{\sigma}_e^{R_i}$ holds (trivial); (vi) $\sigma_{M_i}^{R_i} < \widehat{\sigma}_B^{R_i}$ always holds as $\sigma_{M_i}^{R_i} < \widehat{\sigma}_B^{R_i} \Leftrightarrow 0 < \frac{N}{B_i(1-\beta)} + \frac{\mu(2-N)}{1-\beta} + N + \frac{\beta N\mu}{1-\beta}$, and the RHS of the second inequality is always positive. ■

C Unified versus regional schooling

Proof of Lemma 3 A dominant bourgeoisie prefers R_1 to R_2 whenever $\Pi_B^{R_1} > \Pi_B^{R_2}$, i.e.,

$$I_B^{R_2} - I_B^{R_1} > (1-\beta)\left((1-\delta\mu\sigma)\right)(B_1 - B_2)\right)\frac{M}{2B}.$$

For $\sigma > \widehat{\sigma}_N^{R_1} = \widehat{\sigma}_N^{R_2}$, we have $I_B^{R_2} = I_B^{R_1} = 0$ and R_1 is preferred if $\sigma > \frac{1}{\mu\delta}$, which always holds. The same condition holds for copayment under both systems since $\widetilde{I_B^{R_1}} = \widetilde{I_B^{R_2}}$. If $H^{R_2} < 2$ these are the only relevant comparisons. For $H^{R_2} > 2$ the bourgeoisie fully finances education in each region when $\widehat{\sigma}_B^{R_2} < \sigma < \underline{\sigma_N^{R_2}}$ but the cost is the same in each region and R_1 is preferred as $\sigma > \frac{1}{\mu\delta}$ always holds. For $\widetilde{\sigma}_B^{R_1} < \sigma < \widetilde{\sigma}_B^{R_2}$ again R_1 is preferred since $\widetilde{I_B^{R_1}} < \widehat{I_B^{R_2}} = \frac{M}{2B}$.

We next consider dominant landowners. A dominant landowner prefers R_1 to R_2 whenever $I_N^{R_2} > I_N^{R_1}$. Since $H^{R_1} > H^{R_2}$, the following schooling costs are possible when schooling is implementable in both regions: (i) for $\sigma > \widehat{\sigma}_B^{R_2}$ schooling is free under both systems since $\widehat{\sigma}_B^{R_1} < \widehat{\sigma}_B^{R_2}$, and thus landowners are indifferent; (ii) for $\widehat{\sigma}_B^{R_1} < \sigma < \widehat{\sigma}_B^{R_2}$, $I_N^{R_1} = 0$ and $\widetilde{I_N^{R_2}}$ and hence R_1 is preferred; (iii) for $\widetilde{\sigma}^{R_2} < \sigma < \widehat{\sigma}_B^{R_1}$, the costs are respectively $\widetilde{I_N^{R_1}}$ and $\widetilde{I_N^{R_2}}$. As $\widetilde{I_N^{R_1}} < \widetilde{I_N^{R_2}}$, R_1 is preferred in this case too. ■

Proof of Lemma 4 Dominant landowners prefer U whenever $\Pi_N^U > \Pi_N^{R_i}$, i.e.,

$$(1-\beta)(\delta\sigma - 1)\frac{M}{4} > I_N^U - I_N^{R_i}. \tag{29}$$

The LHS is always positive since the minimum productivity shock for which the landowners are willing to implement U without paying anything is $\underline{\sigma_N^U} = \frac{1}{\delta}$. When landowners can implement U at no cost (i.e. for $\sigma > \widehat{\sigma}_B^U$), U is always preferred as $(1-\beta)(\delta\sigma - 1)\frac{M}{4} > -I_N^{R_i}$ always holds.

In addition, since $\widehat{\sigma}_B^U < \widehat{\sigma}_B^{R_i}$ always holds, schooling at no cost for landowners arises first under U than under R_i. It then only remains to check whether (29) also holds if there is copayment under both systems, i.e. for $\widetilde{I_N^U}$ and $\widetilde{I_N^{R_i}}$, which can only occur if $H^U < 2$.[33] (29) becomes

$$\sigma > \sigma_g = \frac{2 + (1-\beta)\left(N + 2B\left(1-\delta\right) - 2B_i\right)}{(1-\beta)\delta\left(N + 4B\mu - 2\mu B_i\right)}.$$

Since $\widetilde{\sigma}_e^U < \widetilde{\sigma}_e^{R_i}$ is always satisfied, copayment in both systems is only possible for $\sigma > \widetilde{\sigma}_e^{R_i}$. As it is easy to show that $\widetilde{\sigma}_e^{R_i} > \sigma_g$, $\sigma > \sigma_g$ always holds in this case and thus $\Pi_N^U > \Pi_N^{R_i}$ holds also in the case of copayment.

Consider in turn the case of a dominant bourgeoisie. A dominant bourgeoisie prefers U whenever $\Pi_B^U > \Pi_B^{R_i}$, i.e.,

$$(1-\beta)\left(B\delta - B_{-i} + \mu\sigma\delta\left(2B - B_i\right)\right)\frac{M}{2B} > I_B^U - I_B^{R_i}. \tag{30}$$

The LHS is positive for $\sigma > \frac{B_{-i} - B\delta}{\mu\delta(2B - B_i)}$ which is clearly smaller than $\underline{\sigma}_B^{R_i} = \frac{1}{\delta\mu}$. Hence when education is free for the bourgeoisie under both systems, namely for $\sigma > \widehat{\sigma}_N^U = \widehat{\sigma}_{N_i}^{R_i}$, U is preferred. It remains to check what happens under copayment by the bourgeoisie. In this case, U is preferred for $\Pi_B^U(\widetilde{I_B^U}) > \Pi_B^{R_i}(\widetilde{I_B^{R_i}})$ or equivalently for

$$\sigma > \sigma_{gg} = \frac{2 + (1-\beta)\left(N - 2\left(B\delta - B_{-i}\right)\right)}{(1-\beta)\delta\left(N + 2\mu\left(2B - B_i\right)\right)}$$

with $\sigma_{gg} < \widetilde{\sigma}_e^U$ since $\frac{2+(1-\beta)(N-2(B\delta - B_{-i}))}{(1-\beta)\delta(N+2\mu(2B-B_i))} < \frac{2+(1-\beta)(1-B\delta)}{(1-\beta)\delta(N+2B\mu)}$ given that the LHS has a smaller numerator and a bigger denominator than the RHS. As $\widetilde{\sigma}_e^U < \widetilde{\sigma}_e^{R_i}$, U with copayment is always preferred to R_i with copayment by a dominant bourgeoisie. Now if $H^{R_i} > 2$ the bourgeoisie fully finances education in both systems when $\widehat{\sigma}_B^{R_i} < \sigma < \widetilde{\sigma}_B^U$. Then condition (30) becomes

$$\sigma > \frac{1 - (1-\beta)\left(B\delta - B_{-i}\right)}{(1-\beta)\mu\delta\left(2B - B_i\right)}$$

which is smaller than $\widehat{\sigma}_B^{R_i}$ since $-(1-\beta)B_i\left(B\delta - \mu B_i - B_{-i} + 2B\mu\right) < 2B_{-i}$ (notice that the LHS is negative). Hence U is preferred. Now if $H^{R_i} > 2$ then for $\widetilde{\sigma}_B^U < \sigma < \widetilde{\sigma}_e^{R_i}$ the bourgeoisie fully finances regional education but only co-finances it under U, which is clearly better than fully financing it, so U is preferred. ■

D Secession versus unified education

In order to study the incentives of the elite to choose between secession and unified education, we first need to rank the productivity cutoffs under the two systems. This is done in Lemma 5 noting that $H^{S_i} < H^U$.

[33] Observe that $H^{R_2} < H^{R_1} < H^U$. This gives rise to three cases: (i) $H^U < 2$, (ii) $H^{R_1} < 2 < H^U$ and (iii) $H^{R_2} < 2 < H^{R_1}$.

Lemma 5 *1. For $2 > H^U$ then we have either*

- $\underline{\sigma_N} < \widetilde{\sigma}_N^U = \widetilde{\sigma}_B^U < \widetilde{\sigma}_{B_i}^{S_i} = \widetilde{\sigma}_{N_i}^{S_i} < \min\left[\widehat{\sigma}_{N_i}^{S_i} = \widehat{\sigma}_N^U, \widehat{\sigma}_B^U\right] < \widehat{\sigma}_{B_i}^{S_i}$ *or*
- $\underline{\sigma_N} < \widetilde{\sigma}_N^U = \widetilde{\sigma}_B^U < \widehat{\sigma}_B^U < \widetilde{\sigma}_{B_i}^{S_i} = \widetilde{\sigma}_{N_i}^{S_i} < \min[\widehat{\sigma}_{N_i}^{S_i} = \widehat{\sigma}_N^U, \widehat{\sigma}_{B_i}^{S_i}]$

2. For $H^{S_i} < 2 < H^U$ the ranking of the thresholds is $\widehat{\sigma}_B^U < \widetilde{\sigma}_N^U = \widetilde{\sigma}_B^U < \underline{\sigma_N} < \widetilde{\sigma}_{B_i}^{S_i} = \widetilde{\sigma}_{N_i}^{S_i} < \min[\widehat{\sigma}_{N_i}^{S_i} = \widehat{\sigma}_N^U, \widehat{\sigma}_{B_i}^{S_i}]$

3. For $2 < H^{S_i}$ all thresholds but $\widehat{\sigma}_{N_i}^{S_i} = \widehat{\sigma}_N^U$ are smaller than $\underline{\sigma_N}$.

Proof. The three parameter areas follow from Lemma 1 using $H^{S_i} < H^U$. The ordering of the thresholds is based on the following comparisons which mainly use simple algebra (i) $\widehat{\sigma}_N^U = \widehat{\sigma}_{N_i}^{S_i}$; (ii) $\widetilde{\sigma}_N^U = \widetilde{\sigma}_B^U < \widetilde{\sigma}_{B_i}^{S_i} = \widetilde{\sigma}_{N_i}^{S_i}$ always by simple algebra; (iii) $\widehat{\sigma}_B^U < \widehat{\sigma}_{B_i}^{S_i}$ always holds given that $\widehat{\sigma}_B^U < \widehat{\sigma}_{B_i}^{S_i} \Leftrightarrow -(1-\beta)BB_i\,(1+\delta) < 2B_{-i}$ and the second inequality is always verified; (iv) $\widehat{\sigma}_{N_i}^{S_i} > \widehat{\sigma}_{B_i}^{S_i} \Leftrightarrow (1-\beta)NB_i(\mu-1) > N - 2\mu B_i$; (v) $\widehat{\sigma}_{N_i}^{S_i} < \widehat{\sigma}_B^U \Leftrightarrow (1-\beta)BN\,(2\mu-1+\delta) < 2N - 4\mu B$; (vi) $\widehat{\sigma}_B^U > \widetilde{\sigma}_{e_i}^{S_i} \Leftrightarrow 2(N-2\mu B_{-i}) > (1-\beta)B\,((2\mu-1+\delta)\,(2\mu B_i + N) - 4B_i\,(\mu-1)\,\mu)$; (vii) $\widehat{\sigma}_{N_i}^{S_i} > \widetilde{\sigma}_{N_i}^{S_i}$ always; (viii) If $\widehat{\sigma}_{N_i}^{S_i} < \widehat{\sigma}_B^U$ then $\widehat{\sigma}_B^U > \widetilde{\sigma}_{e_i}^{S_i}$ (by point vii). ■

D.1 Same dominant group at the regional and country level

Proof of Proposition 2 Lemma 5 helps us establishing which payment configurations simultaneously arise under U and S_i:

(i) For $\sigma > \widehat{\sigma}_{N_i}^{S_i} = \widehat{\sigma}_N^U$, a dominant bourgeoisie gets schooling for free under both systems. Imposing $I_B^U = I_{B_i}^{S_i} = 0$ in (22), secession is chosen if $0 > (1-\beta)\delta(1+\mu\sigma)\frac{M}{2}$, which never holds.

(ii) Next, whenever $\max(\widetilde{\sigma}_{B_i}^{S_i}, \underline{\sigma_N}) < \sigma < \widehat{\sigma}_{N_i}^{S_i} = \widehat{\sigma}_N^U$, there is copayment under both systems. In that case,

$$\widetilde{I_B^U} - \widetilde{I_{B_i}^{S_i}} = (2 - N(1-\beta)(\delta\sigma - 1))\left(\frac{B_i - B_{-i}}{4BB_i}\right)M.$$

Then, as $(2 - N(1-\beta)(\delta\sigma-1)) > 0$ for $\sigma < \widehat{\sigma}_{N_i}^{S_i} = \widehat{\sigma}_N^U$ we have that $\widetilde{I_B^U} - \widetilde{I_{B_i}^{S_i}} < 0$ for $i = 2$ since $B_1 > B_2$ and hence U is always preferred to S_2 as the payoff (resp. the cost) of schooling is higher (resp. lower) under U. We will now show that condition (22) is also violated with copayment for S_1. Assume by contradiction that condition (22) holds. This would require:

$$\frac{(2 - N(1-\beta)(\delta\sigma-1))\,(B_1 - B_2)}{4BB_1}M > (1+\mu\sigma)(1-\beta)\delta\frac{M}{2}$$

which can be rewritten as

$$\sigma < \sigma_s = \frac{2\,(B_1 - B_2) + (1-\beta)\,(N\,(B_1 - B_2) - 2\delta BB_1)}{(1-\beta)\delta\,(2\mu BB_1 + N\,(B_1 - B_2))}.$$

However, as $\widetilde{\sigma}_{N_1}^{S_1} > \sigma_s$ by simple algebra this is incompatible with the bourgeoisie being willing to pay for the additional cost of education.

(iii) For $\widehat{\sigma}_B^{S_i}, < \sigma < \underline{\sigma}_N$ (which can only hold for $H^{S_i} > 2$), landowners are not willing to contribute to the cost of education, but it is still profitable for the bourgeois under both systems to implement education bearing its full cost. Then, in that case, $\widetilde{I_B^U} = \frac{M}{B}$ and $\widetilde{I_{B_i}^{S_i}} = \frac{M}{2B_i}$ and condition (22) becomes

$$B_i - B_{-i} > (1 + \mu\sigma)(1 - \beta)\delta B B_i,$$

which can never hold for $i = 2$. For $i = 1$ the condition is equivalent to $\sigma < \sigma_{ss} \equiv \frac{B_1 - B_2 - (1-\beta)\delta B B_1}{\mu(1-\beta)\delta B B_1}$ and simple calculus reveals that $\sigma_{ss} < \widehat{\sigma}_{B_1}^{S_1}$, where $\widehat{\sigma}_{B_1}^{S_1}$ is the cutoff for the bourgeoisie to be willing to fully pay education under S_1. Hence condition (22) is violated.

(iv) Finally, for lower values of σ (specifically, $\widetilde{\sigma}_B^U < \sigma < \widetilde{\sigma}_{B_i}^{S_i}$ for $H^{S_i} < H^U < 2$, $\widehat{\sigma}_B^U < \sigma < \widetilde{\sigma}_{B_i}^{S_i}$ for $H^{S_i} < 2 < H^U$, and $\widehat{\sigma}_B^U < \sigma < \widehat{\sigma}_{B_i}^{S_i}$ for $H^{S_i} < H^U < 2$) education is implemented only under U, and thus U is always preferred. ■

Proof of Proposition 3 We need to show that $I_N^U \leq I_{N_i}^{S_i}$. By lemma 5, dominant landowners either get education for free under both systems for $\sigma > \widehat{\sigma}_{B_i}^{S_i}$ or have to copay under S_i but get education for free under U (i.e. $0 = I_N^U \leq \widetilde{I_{N_i}^{S_i}}$) or have to copay under both systems. It remains to prove that $\widetilde{I_N^U} < \widetilde{I_{N_i}^{S_i}}$ which can be rewritten as $2B_{-i}\mu\delta\sigma + B\delta > B_{-i} - B_i$. It is immediate to see that this holds for S_1. For S_2, the required cutoff is $\sigma > \sigma_l = \frac{B_1 - B_2 - B\delta}{2B_1\mu\delta}$ but $\sigma_l < \underline{\sigma_N}$, hence this is always true. ■

D.2 Different dominant group at the regional and country level

D.2.1 Regionally-dominant but countrywide-dominated bourgeoisie

The complete statement of **Proposition 4** is:

The preferences of the bourgeoisie are as follows

1. For $2 < H^{S_i}$ the bourgeoisie always prefers being dominated under unified schooling over being dominant under secession. However, for $\widehat{\sigma}_{B_i}^{S_i} < \underline{\sigma_N}$ unified schooling is not implemented with dominant landowners and the bourgeoisie prefers secession with schooling.

2. For $H^{S_i} < 2 < H^U$ the bourgeoisie always prefers being dominated under unified schooling over being dominant under secession if $B_i > \frac{N}{2\mu}$. If $B_i < \frac{N}{2\mu}$ and $\max[\sigma_{aa}, \widehat{\sigma}_{N_i}^{S_i}] = \widehat{\sigma}_{N_i}^{S_i}$ then the bourgeoisie prefers secession for $\sigma_{aa} < \sigma < \sigma_a$.

3. Let $2 > H^U$.

   (a) If the region-2 bourgeoisie is dominant, then secession will never occur for $B_2 > \frac{N}{2\mu}$. If $B_2 < \frac{N}{2\mu}$, secession might occur for intermediate values of σ. Specifically,

      i. for $\widetilde{\sigma}_{B_2}^{S_2} > \widehat{\sigma}_B^U$ and $\min[\sigma_{aa}, \sigma_a] = \sigma_a$, secession is never preferred

      ii. for $\widetilde{\sigma}_{B_2}^{S_2} > \widehat{\sigma}_B^U$ and $\min[\sigma_{aa}, \sigma_a] = \sigma_{aa}$, secession is preferred for $\sigma_{aa} < \sigma < \sigma_a$.

      iii. for $\widetilde{\sigma}_{B_2}^{S_2} < \widehat{\sigma}_B^U$, secession is preferred for $\widetilde{\sigma}_{B_2}^{S_2} < \sigma < \sigma_a$.

(b) If the region-1 bourgeoisie is dominant, then secession might occur for intermediate values of σ. More precisely

i. if $B_1 > \frac{N}{2\mu}$ then secession is never preferred if $\widetilde{\sigma}^{S_i}_{B_i} > \widehat{\sigma}^{U}_{B}$ but if $\widetilde{\sigma}^{S_i}_{B_i} < \widehat{\sigma}^{U}_{B}$ then secession is preferred for $\widetilde{\sigma}^{S_i}_{B_i} < \sigma < \min[\sigma_a, \sigma_{aa}]$

ii. if $B_1 < \frac{N}{2\mu}$, $\widetilde{\sigma}^{S_1}_{B_1} > \widehat{\sigma}^{U}_{B}$ and $\min[\sigma_{aa}, \sigma_a] = \sigma_a$, then secession is never preferred

iii. if $B_1 < \frac{N}{2\mu}$, $\widetilde{\sigma}^{S_1}_{B_1} > \widehat{\sigma}^{U}_{B}$ and $\min[\sigma_{aa}, \sigma_a] = \sigma_{aa}$, then secession is preferred for $\sigma_{aa} < \sigma < \sigma_a$

iv. if $B_1 < \frac{N}{2\mu}$ and $\widetilde{\sigma}^{S_1}_{B_1} < \widehat{\sigma}^{U}_{B}$, then secession is preferred for $\widetilde{\sigma}^{S_1}_{B_1} < \sigma < \sigma_a$.

where

$$\sigma_a = \frac{2 - (1-\beta)\delta B}{(1-\beta)\mu\delta B} \tag{31}$$

is such that $\Pi^U_B \left(I^U_B = \frac{M}{B}\right) = \Pi^S_{B_i}(I^{S_i}_{B_i} = 0)$ while

$$\sigma_{aa} = \frac{2\left(B_i - B_{-i}\right) - (1-\beta)B\left(N + \delta 2B_i\right)}{(1-\beta)B\delta\left(2\mu B_i - N\right)} \tag{32}$$

is such that $\Pi^U_B \left(I^U_B = \frac{M}{B}\right) = \Pi^S_{B_i}(I^{S_i}_{B_i} = \widetilde{I^{S_i}_{B_i}})$.

Proof of Proposition 4 The bourgeoisie prefers S_i to U whenever condition (22) holds, namely

$$I^U_B - I^{S_i}_{B_i} > (1+\mu\sigma)(1-\beta)\delta\frac{M}{2}. \tag{22}$$

The exact value of I^U_B and $I^{S_i}_{B_i}$ depends on the identity of the dominant group, the size of the shock, and the underlying parameters.

The following payment constellation may occur:

1. $\sigma > \max\left(\widehat{\sigma}^{S_i}_{N_i}, \widehat{\sigma}^{U}_{B}\right)$: from Fig. 3 and 4, if education is implemented under U, the dominated bourgeoisie pays $I^U_B = \frac{M}{B}$ and from Fig. 1 and 2, education for the dominant bourgeoisie is free under secession ($I^{S_i}_{B_i} = 0$). From (22), secession is preferred if and only if $\sigma < \sigma_a$ with σ_a as defined in (31).

2. $\min[\widehat{\sigma}^{S_i}_{N_i}, \widehat{\sigma}^{U}_{B}] < \sigma < \max[\widehat{\sigma}^{S_i}_{N_i}, \widehat{\sigma}^{U}_{B}]$: we have to distinguish two subcases:

   (a) If $\min[\widehat{\sigma}^{S_i}_{N_i}, \widehat{\sigma}^{U}_{B}] = \widehat{\sigma}^{S_i}_{N_i}$, then $I^U_B = \overline{I^U_B}$ and $I^{S_i}_{B_i} = 0$. In this case, S_i is always preferred because condition (22) reduces to $\sigma > \underline{\sigma^{S_i}_{B_i}} = \frac{1}{\mu\delta}$, which is always satisfied as this is the condition for the bourgeoisie to be willing to implement free education under S_i.

   (b) If $\min[\widehat{\sigma}^{S_i}_{N_i}, \widehat{\sigma}^{U}_{B}] = \widehat{\sigma}^{U}_{B}$, then $I^U_B = \frac{M}{B}$ and $I^{S_i}_{B_i} = \widetilde{I^{S_i}_{B_i}}$. The condition that S_i is preferred becomes

$$\sigma < \sigma_{aa_1} = \sigma_{aa} = \frac{2\left(B_i - B_{-i}\right) - (1-\beta)B\left(N + 2\delta B_i\right)}{(1-\beta)B\delta\left(2\mu B_i - N\right)} \text{ for } 2\mu B_i > N \tag{33}$$

$$\sigma > \sigma_{aa_2} = \sigma_{aa} = \frac{(1-\beta)B\,(N+2\delta B_i) - 2\,(B_i - B_{-i})}{(1-\beta)B\delta\,(N-2\mu B_i)} \text{ for } 2\mu B_i < N \tag{34}$$

3. For $\widetilde{\sigma}^{S_i}_{B_i} < \sigma < \max\,[\widehat{\sigma}^{S_i}_{N_i}, \widehat{\sigma}^{U}_{B}]$, $I^U_B = \overline{I^U_B}$ and $I^{S_i}_{B_i} = \widetilde{I^{S_i}_{B_i}}$. Secession is always preferred in this area since condition (22) reduces to $\sigma > \widetilde{\sigma}^{S_i}_{B_i}$ which is the condition for the bourgeoisie to be willing to go for copayment under S_i.

We need to check under which conditions the cutoffs (31), (33) and (34) are relevant cutoffs. Both σ_a and σ_{aa_1} are upper bounds. Therefore σ_a is not relevant if $\sigma_a < \max[\widehat{\sigma}^{S_i}_{N_i}, \widehat{\sigma}^{U}_{B}]$. Similarly, σ_{aa_1} is not relevant for $\sigma_{aa_1} < \max[\widetilde{\sigma}^{S_i}_{B_i}, \widehat{\sigma}^{U}_{B}]$. Since σ_{aa_2} is a lower bound it is not relevant for $\sigma_{aa_2} > \widehat{\sigma}^{S_i}_{N_i}$. Lemma 6 tells us under which conditions these cutoffs are relevant and how they relate to each other and to the different payment areas. ■

Lemma 6 *1.* $\sigma_a > \widehat{\sigma}^{U}_{B}$ *always.*

2. $\sigma_a < \widehat{\sigma}^{S_i}_{N_i} \Leftrightarrow \sigma_{aa_2} > \widehat{\sigma}^{S_i}_{N_i} \Leftrightarrow \sigma_a < \sigma_{aa_2}$.

3. $\sigma_{aa_1} < \widehat{\sigma}^{S_i}_{N_i} \Leftrightarrow \sigma_{aa_1} < \sigma_a \Leftrightarrow \sigma_a < \widehat{\sigma}^{S_i}_{N_i}$.

4. $\widetilde{\sigma}^{S_i}_{B_i} > \widehat{\sigma}^{U}_{B} \Leftrightarrow \sigma_{aa_1} < \widetilde{\sigma}^{S_i}_{B_i} \Leftrightarrow \sigma_{aa_1} < \widehat{\sigma}^{U}_{B}$.

5. For $2\mu B_2 > N$, $\sigma_a < \widehat{\sigma}^{S_2}_{N_2}$ *and* $\sigma_{aa_1} < \underline{\sigma_N}$ *and* $\sigma_{aa_1} < \widetilde{\sigma}^{S_2}_{B_2}$ *always.*

6. $\sigma_{aa_2} < \widetilde{\sigma}^{S_i}_{N_i} \Leftrightarrow \widetilde{\sigma}^{S_i}_{N_i} < \widehat{\sigma}^{U}_{B} \Leftrightarrow \sigma_{aa_2} < \widehat{\sigma}^{U}_{B}$.

7. $\min[\widehat{\sigma}^{S_i}_{N_i}, \widehat{\sigma}^{U}_{B}] = \widehat{\sigma}^{S_i}_{N_i} \Rightarrow \sigma_a > \max[\widehat{\sigma}^{S_i}_{N_i}, \widehat{\sigma}^{U}_{B}]$.

Proof.

1. $\widehat{\sigma}^{U}_{B} = \frac{2-(1-\beta)B(2\mu-1+\delta)}{2\mu(1-\beta)\delta B} + \frac{1}{\delta} < \sigma_a = \frac{2-(1-\beta)B(\delta+\mu)}{(1-\beta)\mu\delta B} + \frac{1}{\delta}$ can be rewritten as $(1-\beta)B\,(\delta+1) < 2$ which is always true.

2. Simple algebra reveals that $\sigma_a < \widehat{\sigma}^{S_i}_{N_i} \Leftrightarrow \sigma_{aa_2} > \widehat{\sigma}^{S_i}_{N_i} \Leftrightarrow$
$$NB(1-\beta)\,(\mu+\delta) > 2(N-\mu B). \tag{35}$$

3. Simple algebra reveals that $\sigma_{aa_1} < \widehat{\sigma}^{S_i}_{N_i} \Leftrightarrow \sigma_{aa_1} < \sigma_a \Leftrightarrow \sigma_a < \widehat{\sigma}^{S_i}_{N_i} \Leftrightarrow$ condition (35) holds.

4. Simple algebra reveals that $\widetilde{\sigma}^{S_i}_{B_i} > \widehat{\sigma}^{U}_{B} \Leftrightarrow \sigma_{aa_1} < \widetilde{\sigma}^{S_i}_{B_i} \Leftrightarrow \sigma_{aa_1} < \widehat{\sigma}^{U}_{B} \Leftrightarrow$
$$(1-\beta)B\,(N\,(2\mu-(1-\delta)) + 2\mu B_i\,(1+\delta)) > 2N - 4\mu B_{-i} \tag{36}$$

5. From the proof of Proposition 4, σ_{aa_1} is the relevant threshold for region-2 secession for $2\mu B_2 > N$. It is easy to see that (35) always holds in this case. Hence by point 3 in this Lemma it follows that $\sigma_a < \widehat{\sigma}^{S}_{N_2}$ always holds for this parameter space. Next $\sigma_{aa_1} < \underline{\sigma_N} = \frac{1}{\delta} \Leftrightarrow (B_i - B_{-i}) < (1-\beta)BB_i\,(\mu+\delta)$ which is always true for $B_i < B_{-i}$ and hence it is always true for secession in region 2. Similarly, $2\mu B_2 > N \Leftrightarrow 2N - 4\mu B_2 < 0$ and thus also $2N - 4\mu B_1 < 0$ given that $B_1 > B_2$. Then, the RHS of (36) is negative for region-2 secession, which from (36) implies that $\sigma_{aa_1} < \widetilde{\sigma}^{S_i}_{B_i}$ holds for region-2 secession.

6. Simple algebra reveals that $\sigma_{aa_2} < \widetilde{\sigma}_{N_i}^{S_i} \Leftrightarrow \widetilde{\sigma}_{N_i}^{S_i} < \widehat{\sigma}_B^U \Leftrightarrow \sigma_{aa_2} < \widehat{\sigma}_B^U \Leftrightarrow$

$$(1-\beta)B\left(N\left(2\mu-(1-\delta)\right)+2\mu B_i\left(1+\delta\right)\right) < 2N - 4\mu B_{-i}. \tag{37}$$

7. If $\min[\widehat{\sigma}_{N_i}^{S_i}, \widehat{\sigma}_B^U] = \widehat{\sigma}_{N_i}^{S_i}$ then $\sigma_a > \max[\widehat{\sigma}_{N_i}^{S_i}, \widehat{\sigma}_B^U]$ since by point 1 $\sigma_a > \widehat{\sigma}_B^U$.

■

We are now set to prove proposition 4. In general, the results follow by combining the parameter restriction and the resulting ranking of the cutoffs with the insights derived from Lemma 6. Here are the details.

1. We look at the parameter area where $2 > H^U$. Given $\widetilde{\sigma}_B^U < \widetilde{\sigma}_{B_i}^{S_i}$ it immediately follows that U is preferred for low σ, namely $\widetilde{\sigma}_N^U = \widetilde{\sigma}_B^U < \sigma < \widetilde{\sigma}_{B_i}^{S_i} = \widetilde{\sigma}_{N_i}^{S_i}$.

   (a) Let $B_i > \frac{N}{2\mu}$. For $i = 2$ by point 5 of Lemma 6 neither σ_{aa_1}nor σ_a are relevant cutoffs and therefore S_i is never preferred. In turn, for $i = 1$ we have to distinguish two further cases: (i) if $\widetilde{\sigma}_{B_i}^{S_i} > \widehat{\sigma}_U^B$ by point 4 of Lemma 6 $\sigma_{aa_1} < \widetilde{\sigma}_{B_i}^{S_i}$ and therefore σ_{aa_1} is not a relevant cutoff. Moreover, since $\widetilde{\sigma}_{B_i}^{S_i} < \widehat{\sigma}_{N_i}^{S_i}$ always by point 3 of Lemma 6, σ_a is not a relevant cutoff either and S_i is never preferred; (ii) If $\widetilde{\sigma}_{B_i}^{S_i} < \widehat{\sigma}_U^B$ by point 4 of Lemma 6 σ_{aa_1} might be a relevant cutoff. By point 3 of Lemma 6 the relevant cutoff is $\min[\sigma_{aa_1}, \sigma_a]$. Notice that σ_a is always relevant by point 7 of Lemma 6 if $\min[\widehat{\sigma}_{N_i}^{S_i}, \widehat{\sigma}_B^U] = \widehat{\sigma}_{N_i}^{S_i}$.

   (b) Let $B_i < \frac{N}{2\mu}$. Again we need to distinguish two cases: (i) If $\widetilde{\sigma}_{B_i}^{S_i} > \widehat{\sigma}_U^B$, then by point 6 of Lemma 6 $\sigma_{aa_2} > \widehat{\sigma}_U^B$. By point 2 of Lemma 6 if $\min[\sigma_{aa_2}, \sigma_a] = \sigma_a$, then S_i is never preferred, but if $\min[\sigma_{aa_2}, \sigma_a] = \sigma_{aa_2}$ then both cutoffs are relevant and S_i is preferred for $\sigma_{aa_2} < \sigma < \sigma_a$. In turn, (ii) if $\widetilde{\sigma}_{B_i}^{S_i} < \widehat{\sigma}_U^B$ by point 6 of Lemma 6 $\sigma_{aa_2} < \widetilde{\sigma}_{B_i}^{S_i}$ and hence the lower bound for secession becomes $\widetilde{\sigma}_{B_i}^{S_i}$. Also $\sigma_{aa_2} < \widehat{\sigma}_U^B$ and by point 2 of Lemma 6 $\sigma_{aa_2} < \widehat{\sigma}_{N_i}^{S_i}$, so σ_a is the relevant upper bound for secession.

2. We now look at the parameter area where $H^{S_i} < 2 < H^U$. In this parameter constellation, the bourgeoisie always fully finances education under U. Since $\widehat{\sigma}_B^U < \widetilde{\sigma}_{B_i}^{S_i}$ by point 4 of lemma 6, the cutoff σ_{aa_1} is never relevant given that $\sigma_{aa_1} < \widetilde{\sigma}_{B_i}^{S_i}$. From point 4, we also know that $\sigma_{aa_1} < \widehat{\sigma}_B^U$ holds. For $2\mu B_i > N$, $\widehat{\sigma}_{N_i}^{S_i} > \widehat{\sigma}_{B_i}^{S_i}$. As $\widehat{\sigma}_{B_i}^{S_i} > \widehat{\sigma}_B^U$, then $\widehat{\sigma}_{N_i}^{S_i} > \widehat{\sigma}_B^U$. Combining this with $\widehat{\sigma}_B^U > \sigma_{aa_1}$, we get that $\widehat{\sigma}_{N_i}^{S_i} > \sigma_{aa_1}$ and thus from point 3, $\widehat{\sigma}_{N_i}^{S_i} > \sigma_a$, and hence σ_a is not a relevant cutoff and U is always preferred. Consider instead now the case with $2\mu B_i < N$. As $\widehat{\sigma}_B^U < \widetilde{\sigma}_{B_i}^{S_i} = \widetilde{\sigma}_N^S$, from point 6 of lemma 6, we know that $\sigma_{aa_2} > \widehat{\sigma}_B^U$ and $\sigma_{aa_2} > \widetilde{\sigma}_{B_i}^{S_i}$. Thus secession with education partly financed by the bourgeoisie is possible if and only if $\sigma_{aa_2} < \widehat{\sigma}_{N_i}^{S_i}$. Note that σ_a is defined as the point of intersection of

$$\Pi_B^U\left(I_B^U = \frac{M}{B}\right) = -\frac{M}{B}\left(1-(1-\beta)\delta B\right) + M(1-\beta)\delta\mu\sigma \tag{38}$$

$$\Pi^S_{B_i}\left(I^{S_i}_{B_i}=0\right)=\delta(1-\beta)\frac{M}{2}+(1-\beta)\frac{M}{2}\mu\sigma. \tag{39}$$

It is easy to check that the intercept of (39) is larger than the intercept of (38) and that its slope with respect to σ is half the slope of (38). In turn, σ_{aa_1} and σ_{aa_2} correspond to the point of intersection of $\Pi^U_B(I^U_B=\frac{M}{B})$ with

$$\Pi^S_{B_i}\left(I^{S_i}_{B_i}=\widetilde{I^{S_i}_{B_i}}\right)=-\frac{M}{4B_i}\left(2-(1-\beta)\left(2B_i\delta-N\right)\right)+\frac{(1-\beta)\left(2\mu B_i+N\right)\delta M}{4B_i}\sigma \tag{40}$$

for respectively $2\mu B_i>N$ and $2\mu B_i<N$. In addition, the slope of (40) is smaller than the slope of (38) if and only if $2\mu B_i>N$. The intercept of (40) is always smaller than the intercept of (38) for S_2. This is also true as long as σ_{aa_2} is positive for S_1.[34] Therefore we have $\sigma_{aa_2}<\sigma_a$ for $2\mu B_2<N$. Then, from point 2, we also have that $\sigma_{aa_2}<\widehat{\sigma}^{S_i}_{N_i}$ and $\widehat{\sigma}^{S_i}_{N_i}<\sigma_a$, which implies that S_2 is chosen by the bourgeoisie for $2\mu B_2<N$ if $\sigma_{aa_2}<\sigma<\sigma_a$.

3. Finally, we study the parameter area $H^{S_i}>2$. In this case $\widehat{\sigma}^U_B<\underline{\sigma_N}<\widehat{\sigma}^{S_i}_{N_i}$ always holds and hence the bourgeoisie always fully finances education under U. We also have $\widetilde{\sigma}^{S_i}_{B_i}<\underline{\sigma_N}$, so that the bourgeoisie is always willing to go for copayment for $\underline{\sigma_N}<\sigma<\widehat{\sigma}^{S_i}_{N_i}$ and will get education for free for $\sigma>\widehat{\sigma}^{S_i}_{N_i}$. In this parameter area, $\widetilde{\sigma}^{S_i}_{B_i}>\widehat{\sigma}^U_B$ always holds. Then by point 4 (rep. point 3) of lemma 6 $\sigma_{aa_1}<\widetilde{\sigma}^{S_i}_{B_i}$ (resp. $\sigma_a<\widehat{\sigma}^{S_i}_{N_i}$) which implies that σ_{aa_1} (resp. σ_a) is never a relevant cutoff and that U is always preferred. Now, by point 6 of lemma 6 $\sigma_{aa_2}>\widehat{\sigma}^U_B$. Can it be the case that secession is preferred for $\sigma_{aa_2}<\sigma<\sigma_a$? From (34), the cutoff σ_{aa_2} might only be relevant if $2\mu B_i<N$. Note that this can be rewritten as $2\mu B_i+B<1$. The parameter area we are studying requires $H^{S_i}=2(1-\beta)(\mu-1)B_i>2$. Combining both conditions requires $2\mu B_i+B<(1-\beta)(\mu-1)B_i$ which is equivalent to $(2\mu-(1-\beta)(\mu-1))\,B_i+B<0$ which is clearly false. Hence σ_{aa_2} cannot be a relevant cutoff when $H^{S_i}>2$. So U is always preferred when it is implementable. However, education under S_i is implemented by the bourgeoisie before education under U, namely for $\widehat{\sigma}^{S_i}_{B_i}<\sigma<\underline{\sigma_N}$. For these parameter values, the dominant bourgeoisie in region i prefers S_i to being dominated by landowners with no schooling under U. ■

D.2.2 Regionally-dominant but countrywide-dominated landowners

Proof of Proposition 5 If landowners are dominant in region i but dominated at the country level, they prefer secession whenever $I^U_N>I^S_{N_i}$. The following educational costs are possible

- For $2>H^U$, there are two possible rankings of the cutoffs:
  1. $\widetilde{\sigma}^U_N<\widetilde{\sigma}^{S_i}_{N_i}<\widehat{\sigma}^U_N<\widehat{\sigma}^{S_i}_{B_i}$ (Lemma 5). For $\sigma>\widehat{\sigma}^{S_i}_{B_i}$ education is free under S_i but has to be paid fully under U, so S_i is preferred. For $\widehat{\sigma}^U_N<\sigma<\widehat{\sigma}^{S_i}_{B_i}$, under U, N landowners pay full education costs for M mass members, while under S_i there is copayment and thus $N/2$ landowners pay less than the full cost for $M/2$ mass members implying that $I^S_{N_i}<I^U_N$ and thus S_i is preferred. Next, for $\widetilde{\sigma}^{S_i}_{N_i}<\sigma<\widehat{\sigma}^U_N$ landowners pay their

[34] But the opposite holds if σ_{aa_2} is negative.

maximum willingness $\overline{I_N^U} = \frac{(\delta\sigma-1)(1-\beta)M}{2}$ under U and copay $\widetilde{I_{N_i}^{S_i}} = \frac{1-B_i(1-\beta)(\delta\mu\sigma-1)}{N}M$ under S_i. As it can be shown that $I_N^U > I_{N_i}^S \Leftrightarrow \sigma > \widetilde{\sigma}_{N_i}^{S_i}$, S_i is also preferred by the landowners whenever $\sigma > \widetilde{\sigma}_{N_i}^{S_i}$, i.e., whenever it is implementable. For $\widetilde{\sigma}_N^U < \sigma < \widetilde{\sigma}_{N_i}^{S_i}$, education is only implemented under U, and thus the landowners prefer U in that case.

2. $\widetilde{\sigma}_N^U < \widetilde{\sigma}_{N_i}^{S_i} < \widehat{\sigma}_{B_i}^{S_i} < \widehat{\sigma}_N^U$ (Lemma 5). For $\sigma > \widehat{\sigma}_{B_i}^{S_i}$, education is free under S_i but the landowners either pay their maximal willingness or the entire education under U, so S_i is preferred. For $\widetilde{\sigma}_{N_i}^{S_i} < \sigma < \widehat{\sigma}_N^U$, landowners pay their maximum willingness $\overline{I_N^U} = \frac{(\delta\sigma-1)(1-\beta)M}{2}$ under U and copay $\widetilde{I_{N_i}^{S_i}} = \frac{1-B_i(1-\beta)(\delta\mu\sigma-1)}{N}M$ under S_i and $I_N^U > I_{N_i}^S$ whenever $\sigma > \widetilde{\sigma}_{N_i}^{S_i}$. Finally, for $\widetilde{\sigma}_N^U < \sigma < \widetilde{\sigma}_{N_i}^{S_i}$, education is only implemented under U, and thus the landowners prefer U in that case.

- For $H^{S_i} < 2 < H^U$, $\widetilde{\sigma}_N^U < \underline{\sigma_N} < \widetilde{\sigma}_{N_i}^{S_i} < \min[\widehat{\sigma}_N^U, \widehat{\sigma}_{B_i}^{S_i}]$ holds (Lemma 5). If $\min[\widehat{\sigma}_N^U, \widehat{\sigma}_{B_i}^{S_i}] = \widehat{\sigma}_N^U$, then we have the same cases as for $2 > H^U$ and secession is always preferred (for $\sigma > \widetilde{\sigma}_{N_i}^{S_i}$) when it is implementable. If $\min[\widehat{\sigma}_N^U, \widehat{\sigma}_{B_i}^{S_i}] = \widehat{\sigma}_{B_i}^{S_i}$ then for $\sigma > \widehat{\sigma}_{B_i}^{S_i}$ education is free under S_i and the landowners have to pay their maximal willingness (and later even the entire education) under U, so S_i is preferred. For $\widetilde{\sigma}_{N_i}^{S_i} < \sigma < \widehat{\sigma}_{B_i}^{S_i}$, there is copayment under S_i while the landowners have to pay their maximal willingness under U, and we have shown that S_i is preferred for $\sigma > \widetilde{\sigma}_{N_i}^{S_i}$. Hence also here S_i is always preferred when it is implementable.

- For $H^{S_i} > 2$, landowners get education under S_i for free, so it is always preferred when it is implementable. ■

E Welfare analysis

E.1 Proof of Proposition 7

R_1 is preferred to U by the social planner if and only if $1 + B_2 + \frac{N}{2} + \delta\left(1 + \sigma\left(\mu B_1 + \frac{N}{2}\right)\right) > 2\delta\left((1+\sigma)\frac{N}{2} + (1+\mu\sigma)B\right)$, which is equivalent to $\sigma < \sigma_z \equiv \frac{1+\left(\frac{N}{2}+B_2\right)-\delta B}{\delta\left(\frac{N}{2}+\mu(B_2+B)\right)}$. However, it is easy to show that $\sigma_z < \sigma_W^{R_1S_1}$ which means that R_1 is preferred to U only when the social planner prefers no education to R_1, and thus the implementation of R_1 is never a first best outcome. The same applies to R_2 as $\sigma_W^{R_1S_1} < \sigma_W^{R_2S_2}$. Finally, regional schooling with schooling in both regions is dominated by U given that the cost of both systems is the same and only U generates cross-regional matches. ■

E.2 Proof of Proposition 6

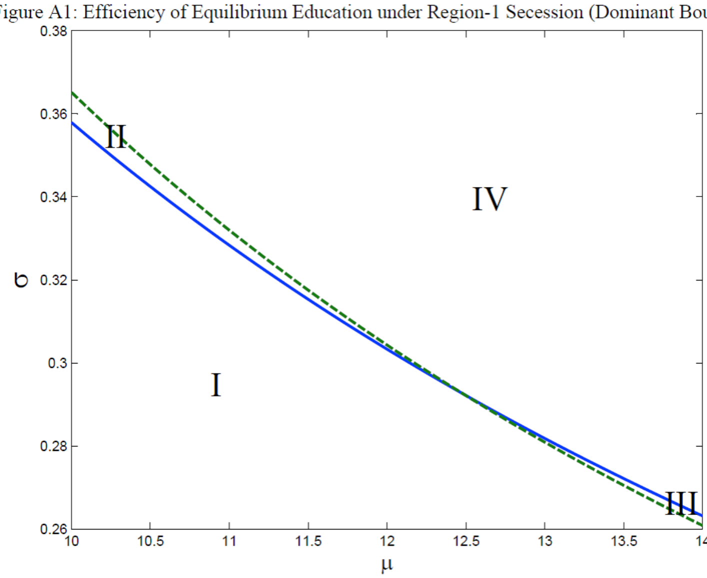
We first show that when $H^U < 2$, U results in undereducation for $\sigma_W^U \leq \sigma \leq \widetilde{\sigma}_e^U$. Indeed, for $H^U < 2$, the threshold under U is always $\widetilde{\sigma}_e^U$ no matter the identity of the dominant group, and it is easy to show that $\sigma_W^U < \widetilde{\sigma}_e^U$ always holds. Next we consider the case where

$H^U > 2$. When the landowners are dominant, the relevant cutoff for education is $\underline{\sigma_N}$. As $H^U \equiv (1-\beta)\,B\,(2\mu+\delta-1) > 2$ implies that $B\,(2\mu+\delta-1) > 2$ and it can be shown that $\sigma_W^U < \frac{1}{\delta} = \underline{\sigma_N} \Leftrightarrow B\,(2\mu+\delta-1) > 2$, then $\sigma_W^U < \underline{\sigma_N}$ holds in this case, and there is undereducation for $\sigma_W^U \leq \sigma \leq \underline{\sigma_N}$. If instead the bourgeoisie is dominant, the relevant cutoff is $\widehat{\sigma}_B^U$. It is easy to show that $\sigma_W^U > \widehat{\sigma}_B^U \Leftrightarrow H^U > 2 + \frac{4\beta B\mu}{N}$. Thus, for $H^U < 2 + \frac{4\beta B\mu}{N}$, $\sigma_W^U < \widehat{\sigma}_B^U$ and there is undereducation for $\sigma_W^U < \sigma < \widehat{\sigma}_B^U$, and instead for $H^U < 2 + \frac{4\beta B\mu}{N}$, $\sigma_W^U > \widehat{\sigma}_B^U$ holds and there is overeducation for $\widehat{\sigma}_B^U < \sigma < \sigma_W^U$. ■

E.3 Efficiency of Region-1 Secession

Figure A1 studies the optimality of schooling under S_1 when the bourgeoisie is dominant and $\beta = 0.1$, $\delta = 0.95$, $B_1 = 0.45$, $B = 0.5$, and $M = 2$ The dashed line represents the threshold for the implementation of education under S_1 for a dominant bourgeoisie (given by $\widehat{\sigma}_B^{S_1}$ as for these parameter values $H^{S_1} > 2$), while the solid line represents the socially-efficient productivity threshold ($\sigma_W^{R_1 S_1}$).

Figure A1: Efficiency of Equilibrium Education under Region-1 Secession (Dominant Bourgeoisie)



As for U, the efficient no-education (resp. education) decision is taken in region I (resp. region IV), while there is undereducation in region II and overeducation in region III.

E.4 Proof of Proposition 8

From Proposition 4, U cannot be implemented but S_i can for $H^{S_i} > 2$ and $\widehat{\sigma}_{B_i}^{S_i} < \sigma < \underline{\sigma_N}$. As $H^{S_i} > 2 \Leftrightarrow \widehat{\sigma}_{B_i}^{S_i} < \underline{\sigma_N} \Leftrightarrow (1-\beta)(\mu-1)B_i > 1$, then we know that $\sigma_W^{R_i S_i} < \underline{\sigma_N} \Leftrightarrow (\mu-1)B_i > 1$ necessarily holds here. As S_i is preferred to no education whenever $\sigma > \sigma_W^{R_i S_i}$, S_i is optimal and implementable for $\max[\sigma_W^{R_i S_i}, \widehat{\sigma}_{B_i}^{S_i}] < \sigma < \underline{\sigma_N}$. ■

E.5 Numerical examples of overprovision and underprovision of secession by a regional dominant but countrywide dominated bourgeoisie

Example 1 *For* $\beta = 0.1$, $N = 0.3, B_1 = 0.5$, $\delta = 0.9$ *and* $\mu = 3.5$ *we have* $\sigma_W^{R_i S_i} = 0.964912 < \widehat{\sigma}_{B_i}^{S_i} = 1,02293 < \underline{\sigma_N} = 1.1111$. *Implemented secession is socially optimal for* $\widehat{\sigma}_{B_i}^{S_i} < \sigma < \underline{\sigma_N}$ *but there is too little secession for* $\sigma_W^{R_i S_i} < \sigma < \widehat{\sigma}_{B_i}^{S_i}$.

Example 2 *For* $\beta = 0.05$, $N = 0.45, B_1 = 0.5$, $\delta = 0.95$ *and* $\mu = 5$ *we have* $\widehat{\sigma}_{B_i}^{S_i} = 0.65374 < \sigma_W^{R_i S_i} = 0.666345 < \underline{\sigma_N} = 1.05263$. *Implemented secession is socially optimal for* $\sigma_W^{R_i S_i} < \sigma < \underline{\sigma_N}$ *but there is too much secession for* $\widehat{\sigma}_{B_i}^{S_i} < \sigma < \sigma_W^{R_i S_i}$.

References

[1] Acemoglu, Daron and James Robinson (2001) "A Theory of Political Transitions", *American Economic Review*, 2001, 91:4, 938–963.

[2] Ager, Philipp (2013) "The Persistence of De Facto Power: Elites and Economic Development in the U.S. South, 1840-1960", *EHES Working Papers in Economic History*, No. 38.

[3] Aghion, Philippe, Xavier Jaravel, Torsten Persson, and Dorothée Rouzet (2013) "Education and Military Rivalry", Mimeo.

[4] Alesina, Alberto and Bryony Reich (2013) "Nation building", *NBER Working Paper* No. 18839.

[5] Armstrong, John (1982), *Nations Before Nationalism*. Chapel Hill, University of North Carolina Press, 1982.

[6] Aspachs-Bracons, Oriol, Irma Clots-Figueras, Joan Costa-Font and Paolo Masella (2008) "Compulsory Language Educational Policies and Identity Formation" *Journal of the European Economic Association*, 6 (2-3), 434-444, April-May 2008.

[7] Balcells, Laia (2013) "Mass Schooling and Catalan Nationalism", *Nationalism and Ethnic Politics*, 19:4, 467-486.

[8] Balfour, Sebastian (1995) "Riot, Regeneration, and Reaction: Spain in the Aftermath of the 1898 Disaster", *The Historical Journal*, Vol. 38, No. 2 (jun., 1995), pp. 405-423.

[9] Becker, Sasha O., Erik Hornung and Ludger Woessmann (2011) "Education and Catch-up in the Industrial Revolution", *American Economic Journal: Macroeconomics* 3: 92-126

[10] Bowles, Samuel (1978) "Capitalist Development and Educational Structure", *World Development,* vol. 6, pp. 783-796.

[11] Breuilly, John (1993) *Nationalism and the State*. Manchester University Press. Second edition.

[12] Breuilly, John (2006) Introduction, in *Nations and Nationalism*, second edition, Blackwell Publishing.

[13] Calvet, Louis-Jean (2002) *Linguistique et colonialisme*. Editions Payot et Rivages.

[14] Carr, Raymond (1980) "The Regional Problem in Spain", in *Federalism: History and Current Significance of a Form of Government*, Johan Chirstiaan Boogman and G.N. van der Plaat (eds), The Hague: Martinus Nijhoff, 1980.

[15] Chervel, André (1992) *L'enseignement du français à l'école primaire*. Textes officiels, Tome 1, 1791–1879. Economica.

[16] Clots-Figueras, Irma and Paolo Masella (2013) "Education, Language, and Identity", *Economic Journal*, 123 (570), F332-357.

[17] Crayen, Dorothee and Joerg Baten (2010) "New Evidence and New Methods to Measure Human Capital Inequality Before and During the Industrial Revolution: France and the US in the Seventeenth to Nineteenth Centuries", *Economic History Review*, 63, 2 (2010), pp. 452-478.

[18] Darden, Keith and Anna Grzymala-Busse (2006) "The Great Divide: Literacy, Nationalism, and the Communist Collapse", *World Politics*, Vol. 59, No. 1 (Oct., 2006), pp. 83-115.

[19] Díez Medrano, Juan (1994) "Patterns of Development and Nationalism: Basque and Catalan Nationalism before the Spanish Civil War", *Theory and Society,* vol. 23, No. 4 (Aug., 1994), pp. 541-569.

[20] Enrlich, Charles E. (1998) "La Lliga Regionalista and the Catalan Industrial Bourgeoisie", Journal of Contemporary History, Vol. 33, No. 3 (Jul., 1998), pp. 399-417.

[21] Galor, Oded and Omer Moav (2006) "Das Human-Kapital: A Theory of Demis of the Class Structure", *The Review of Economic Studies,* 73(1): 85-117

[22] Galor, Oded, Omer Moav, and Dietrich Vollrath (2009) "Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence", *Review of Economic Studies,* vol. 76, No. 1 (Jan. 2009), pp. 143-179.

[23] Gellner, Ernst (1964) *Thought and Change*. Weidenfeld and Nicolson, London.

[24] Gellner, Ernst (1983, 2006) *Nations and Nationalism*. First edition: Cornell University Press. Second edition: Blackwell.

[25] Harrison, Joseph (1974) "Catalan Business and the Loss of Cuba, 1898-1914" *The Economic History Review*, New Series, Vol. 27 No. 3 (Aug., 1974), pp. 431-441.

[26] Harrison, Joseph (1990) "The Catalan Industrial Elite, 1898-1923", in *Elites and Power in Twentieth-Century Spain,* Frances Lannon and Paul Preston (eds.), Clarendon Press, Oxford.

[27] Hastings, Adrian (1997) *The Construction of Nationhood: Ethnicity, Religion, and Nationalism*, Cambridge University Press.

[28] Hippe, Ralph and Jörg Baten (2013) "'Keep Them Ignorant.' Did Inequality in Land Distribution Delay Regional Numeracy Development", Chap. 6, Ralph Hippe's PhD Dissertation, University of Tübingen and University of Strasbourg.

[29] Keating, Michael (1993) "Peripheral Nationalism and State Response", in *The Politics of Ethnic Conflict Regulation*, John McGarry and Brendan O'Leary (eds.), Routledge.

[30] Kroneberg, Clemens and Andreas Wimmer (2012) "Struggling over the Boundaries of Belonging: a Formal Model of Nation Building, Ethnic Closure, and Populism", *American Journal of Sociology*, Vol. 118, No. 1 (July 2012), pp. 176-230.

[31] INE (1991) *Censo de 1787 Floridablanca*. Tomo 6. Artes Gráficas Sorual. Madrid.

[32] Lachapelle, Georges (1928) "Elections législatives 22-29 Avril 1928, Résultats officiels". Librairie Georges Roustan.

[33] Lindert, Peter H. (2004) *Growing Public 1: Social Spending and Economic Growth since the Eighteenth Century*. Cambridge University Press.

[34] Linz, Juan (1974) "Early State-Building and Late Peripheral Nationalisms against the State: the Case of Spain" in *Building States and Nations, Analyses by Regions,* S.N. Eisenstadt and Stein Rokkan (eds.), Sage Publications.

[35] Linz, Juan (1975) "Politics in a Multi-lingual Society with a Dominant World Language: the Case of Spain" in *Multilingual Political Systems: Problems and Solutions.* Jean-Guy Savard and Richard Vigneault (eds.), Les Presses de l'Université Laval.

[36] Maddison, Angus (2003) *The World Economy, Historical Statistics*. Development Centre Studies, OECD.

[37] Nadal, Jordi (1973) "The failure of the industrial revolution in Spain, 1830-1914", in C. Cipolla, ed., *The Fontana economic history of Europe, iv, The emergence of industrial societies,* pp. 532-626.

[38] Nuhoğlu Soysal, Yasemin and David Strang "Construction of the First Mass Education Systems in Nineteenth-Century Europe, *Sociology of Education*, vol. 62, No. 4 (Oct., 1989), pp. 277-288.

[39] Ortega, Javier and Thomas Tangerås (2008) "Unilingual versus Bilingual Education: A Political Economy Analysis", *Journal of the European Economic Association*, 6(5):1078-1108. September 2008.

[40] Price, Roger (2004) *People and Politics in France, 1848-1870.* Cambridge University Press.

[41] Roeder, Philip G. (2007) *Where Nation-States Come From: Institutional Change in the Age of Nationalism.* Princeton University Press.

[42] Shubert, Adrian (1990) *A Social History of Modern Spain,* Unwin Hyman, London.

[43] Smith, Anthony D. (2000) *The Nation in History, Historiographical Debates about Ethnicity and Nationalism.* The Menahem Stern Jerusalem Lectures. University Press of New England.

[44] Solé Tura, Jordi (1989) "The Spanish Case: Remarks on the General Theories of Nationalism", *International Political Science Review/Revue Internationale de Science Politique*, Vol. 10, No. 3 (Jul., 1989), pp. 183-189.

[45] Squicciarini, Mara P. and Nico Voigtländer (2014) "Human Capital and Industrialization: Evidence from the Age of Enlightenment", Mimeo.

[46] Tortella, Gabriel (1994) "Patterns of Economic Retardation and Recovery in South-Western Europe in the Nineteenth and Twentieth Centuries", *The Economic History Review*, New Series, Vol. 47, No. 1 (Feb., 1994), pp. 1-21.

[47] Tortella, Gabriel (2000) *The Development of Modern Spain. An Economic History of the Nineteenth and Twentieth Centuries.* Harvard University Press, Cambridge, Mass.

[48] Tusell, Javier (1982) "Las Constituyentes de 1931: unas elecciones de transición". Centro de Investigaciones Sociológicas. Madrid.

[49] Weber, Eugen (1976). *Peasants into Frenchmen. The Modernization of Rural France*, 1870-1914. Stanford University Press.

[50] West, E.G. (1978) "Literacy and the Industrial Revolution", *The Economic History Review* 31(3): 369-383

[51] Wimmer, Andreas and Yuval Feinstein (2010), "The Rise of the Nation-State across the World, 1816 to 2001", *American Sociological Review* 75 (5), 764-790.