



City Research Online

City St George's, University of London

Citation: Ortega, J. & Tangeraas, T. (2008). Unilingual versus Bilingual Education: a Political Economy Analysis. *Journal of the European Economic Association*, 6(5), pp. 1078-1108. doi: 10.1162/jeea.2008.6.5.1078

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/3593/>

Link to published version: <https://doi.org/10.1162/jeea.2008.6.5.1078>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Unilingual versus bilingual education: A political economy analysis*

Javier Ortega

Toulouse School of Economics (GREMAQ, IDEI) and CEP (LSE)

Thomas P. Tangerås

Research Institute of Industrial Economics (IFN), Stockholm

September 2007

forthcoming in the *Journal of the European Economic Association*

Abstract

We consider an economy with two language groups, where only agents who share a language can produce together. Schooling enhances the productivity of students. Individuals attending a unilingual school end up speaking the language of instruction only, while bilingual schools render individuals bilingual at the same cost. The politically dominant group (not necessarily the majority) chooses the type(s) of schools accessible to each language group, and then individuals decide whether to attend school. We show that the dominant either choose laissez-faire or restrict access to schools in the language of the dominated. Instead, the dominated favour the use of their own language. Thus, while agents do not derive utility from speaking their mother tongue, language conflicts of the expected type endogenously arise. Democracy (majority rule) always leads to the implementation of a socially optimal education system, while restrictions to the use of the language of the dominated are implemented too often under minority rule. The model is consistent with evidence from Belgium, France, and Finland.

JEL classification : I2, J15.

Keywords : education, language policies, skills, minorities

*We would like to thank Alain Alcouffe, Raquel Fernández, Jean Fraysse, Patrick González, Barbara Petrongolo, Anne Ruiz-Gazen, François Salanié, Jean Tirole, Etienne Wasmer, and seminar participants in Valencia, Toulouse, T2M (Paris), Universidad Pública de Navarra, IFN, EALE (Paris), LSE, SED (Paris), Helsinki, Tinbergen Institute, UQAM, EEA (Madrid), ESSLE-CEPR, Pompeu Fabra, Carlos III, Rennes, Tilburg, Brunel, the ECORE Conference “Challenges of Multilingual Societies”, and FEDEA for interesting comments and helpful discussions. Päivi Erkkilä, Marjatta Latvus, Doh-Shin Jeon, Mika Helander, Eric Malin, and Kenneth Snellman helped us gather information. We are also grateful to Christina Lönnblad for her editorial assistance. Tangerås gratefully acknowledges funding from the Jan Walland and Tom Hedelius Research Foundation. Ortega is also affiliated to CEPR, FEDEA, and IZA. Addresses for correspondence: javier.ortega@TSE-fr.eu; thomas.tangeras@ifn.se.

1 Introduction

In 2000, half of the world’s countries had at least one language minority constituting more than 10% of their population. This language diversity has recently brought language policies to the forefront of political debate in such countries as Malaysia, the ex-Soviet States, Spain, Belgium, and the United States. As stressed by sociolinguists, one crucial component behind language shifts across generations is the choice of the language(s) of instruction in school. For example, Fishman (1977, p. 116) argues that “for language spread, schools have long been the major formal (organized) mechanisms involved.” In other words, languages that are not given the status of medium of instruction in school tend to be replaced by languages that are.

The cases of France and Finland provide two illustrations of the importance of language policies for language development. In the late 18th century, about 60% of those living in France did not actually speak French (Grégoire 1794). Nowadays, nearly everybody there speaks French, and other languages are spoken by only 5% of the population (Encyclopaedia Britannica 2003). Instrumental in this development was the implementation of a unilingual education system from the 1880s, that established French as the sole language of instruction in school. At the other end of the spectrum, the bilingual Finnish–Swedish education system, implemented upon Finland’s independence in 1917, has been one of the factors explaining the relatively good shape of Swedish in contemporary Finland, with the number of native Swedish speakers roughly unchanged since 1920.¹

Given the importance of language of instruction, we set up a model to explain why some multilingual countries choose unilingual education while others maintain language diversity. We consider an economy with two language groups whose members are initially unable to communicate. Value is generated from bilateral production after schooling among agents speaking the same language. Schooling is a “bundle” because it enhances the productivity (or earnings) of students and can also modify their language endowment.² Schools can be bilingual or unilingual in either language. Individuals attending a unilingual school end up speaking the language of instruction only, whereas bilingual schools make individuals bilingual at the same cost. Thus, bilingual schools have a technological advantage over unilingual schools.

The politically “dominant” language group (not necessarily the majority) decides first the type(s) of schools accessible to each language group (the “education system”), and then individuals choose whether or not to attend school. The interaction among individual school attendance decisions is generated by the network externalities arising from the requirement that production partners

¹In 1920 there were 314,000 native Swedish speakers (McRae 1997) and in 2000 there were 293,000 (Encyclopaedia Britannica 2003). In relative terms, the number of Swedish speakers has declined from 11% to 5.9% of the total Finnish population.

²The positive effect of education on earnings is a well-established fact in the literature (see e.g. Card 1999), and language of instruction in school is an important factor behind language shift (see e.g. Fishman 1977 or Hagège 1996). The bundling assumption implies that we assume away the possibility that individuals go to schools that exclusively provide language training. This is because we are interested here in the choice of the language of instruction and not in language training in general.

speak a common language.

Although the number of potential education systems is large, we show that no system yields a higher utility to the dominant group than *laissez-faire* (free choice of school) or one of the following two systems that restrict the use of the language of the dominated group: (i) a unilingual system where only the mother tongue of the dominant group is allowed in education; and (ii) an (asymmetric) bilingual system, with bilingual schools for the dominated group and unilingual schools in their own mother tongue for the dominant group.

Under *laissez-faire*, each individual undertaking education chooses to attend a bilingual school, which opens up for more production possibilities than unilingual schools at no additional cost. The attractiveness of *laissez-faire* to the dominant group then simply stems from the exclusive use of the superior schooling technology. However, the dominant may prefer restricting the use of the language of the dominated in order to foster the schooling incentives of the dominated and to have them carry the cost of intergroup communication. Under asymmetric bilingualism, schooling incentives for the dominated are higher than under *laissez-faire* because the dominant legally restrict themselves from learning the language of the dominated. Then, in contrast to *laissez-faire* (where the dominant could become bilingual), the dominated here always gain production partners when attending school. Under unilingualism, these incentives can be even stronger because an uneducated member of the dominated group loses the ability to communicate with the members of her own group who undertake education (a “bandwagon” effect).

All political tension arising in equilibrium is of the expected type—namely, situations in which the dominant want to restrict the use of the language of the dominated while the dominated prefer a system in which their native language is also a language of instruction. This is an interesting result in that it does not rely on any direct utility enjoyed by the agents from speaking their own native language. The dominated want their language to be used in schools not because they “like it” but rather because abandoning it would systematically force them to carry the cost of intergroup communication.

We determine the socially optimal education system. When a benevolent planner can choose the education level of each individual, *laissez-faire* is always optimal owing to the technological superiority of bilingual schools. If the central planner can choose the education system but choice of school attendance remains in the hands of the individuals, then *laissez-faire* is no longer necessarily optimal, since bilingualism or unilingualism may be more effective at inducing higher education levels in equilibrium.

Next we address the issue of failure in political decision making, that is, we analyse the circumstances, if any, under which the political decision process leads to adoption of the “wrong” education system. From a welfare viewpoint, cost-efficient communication implies that the minority learn the majority language, while no system yields a higher utility to the dominant than *laissez-faire* or a restriction to the use of the language of the dominated. Because the dominant group and the majority are the same under majority rule (democracy), this system is shown always to lead to the adoption of a socially optimal decentralised system. If instead the dominant group is a minority (autocracy), then restrictions to the use of the dominated group language are too often

implemented.

The basic model is then extended to consider education subsidies, mother tongue persistence in unilingual schools, cross-border spillovers, and a higher cost of bilingual schools.

Empirically, our model predicts that the size of the language majority may not be the most relevant factor for understanding the choice of education system. Using regional data for 1860s France, we show that the proportion of French-speaking schools is unrelated to the proportion of local French speakers. In addition, the implementation of French unilingualism in 19th-century Belgium is consistent with an elite-driven choice in our model, and the open economy version of our model predicts the unanimous choice of bilingual schools in 1920s Finland.

Our model is related to the growing literature on language adoption and in particular to Lazear (1999), Church and King (1993), and John and Yi (2001).³ As in these three papers, agents in our model choose whether to make a costly investment in learning a language that can be used in trade or production with other agents. However, we differ from these papers by considering an investment decision that ties skill acquisition and language acquisition. In Lazear (1999), agents behave competitively, but in our model—just as in Church and King (1993) and John and Yi (2001)—the investment decision is strategic and the equilibrium outcome depends on a network externality.⁴ Our paper differs from the two latter works in that we endogeneise the type of network externalities under consideration, since it depends on the choice of education system. Another difference is that our explanation of language shift is based on the choice of schooling institutions, whereas John and Yi (2001) provides an explanation based on geography and on intergenerational language transmission. Finally, a further contribution of our model is the derivation of language conflict or consensus as an equilibrium outcome.

2 The model

Consider a country inhabited by a continuum of individuals, normalised to unity. There are two language groups in the country, $L = \{M, N\}$, of respective sizes $p_l = \{p_m, p_n\}$. We also denote by $-L$ the language group other than L . Political power is in the hands of language group M whether or not $p_m \geq 1/2$. For this reason, the M s are also referred to throughout the paper as the “dominant” group and the N s as the “dominated” group.⁵ Initially, the M s

³There are, of course, other papers studying language. Lang (1986) proposes a language theory of discrimination. Pool (1991) and Laitin (1994) analyse the choice of an official language in multilingual countries. Mélitz (2002) shows that sharing a common language promotes international trade. In addition, there is a large literature on language proficiency and earnings (see e.g. Chiswick and Miller 1995) and a new literature on the possible linguistic organisation of the European Union (see Ginsburgh, Ortuño-Ortín, and Weber 2005).

⁴The economics of networks has been extensively studied in the industrial organisation literature; see Farrell and Klemperer (2007) for a recent survey. Research along this line has generally focused on the problem of adaption and coordination from the perspective of profit-maximising firms.

⁵The assignment of the dominant role to the M group is without loss of generality. We do not explain here the reasons why one group becomes the politically dominant group. This

speak *Mish* and the *Ns* speak *Nish*. We assume that communication between two agents is possible if and only if they speak a common language.

Value is created through bilateral production between individuals. Each individual has the opportunity to produce once with every other individual. Bilateral production occurs if and only if the two partners are able to communicate; if they cannot communicate, the value of production is equal to zero.

2.1 Schools

Individuals choose whether or not to attend school. An individual who undertakes education becomes skilled and produces $1 + \sigma$ (with $\sigma > 0$) upon meeting any agent with whom she is able to communicate. An uneducated individual produces 1 with any partner speaking the same language.

Schooling also involves language training, depending on the school characteristics. Schools can be bilingual, *N* unilingual, or *M* unilingual. The personal cost c of undertaking education is assumed to be constant across the population and independent of the school type.

Anyone attending a bilingual school becomes bilingual. The expected utility of attending a bilingual school is

$$U^b = -c + 1 + \sigma; \quad (1)$$

that is, an individual who attends school pays c , becomes skilled and ends up speaking both languages, and thus produces $1 + \sigma$ with every individual in the economy.

In an *L* unilingual school, *Lish* is the unique language of instruction. We assume that anyone undertaking unilingual *Lish* education ends up speaking *Lish* only. Thus, for instance, a native *Nish* speaker who attends a unilingual *M* school learns *Mish* and loses her initial language.⁶ Let μ_l^{lu} (resp., μ_l^{-lu}) be the fraction of native *Lish* speakers who attend an *L* unilingual (resp., $-L$ unilingual) school. Similarly, μ_l^b denotes the proportion of native *Lish* speakers who attend a bilingual school. Finally, μ is the vector of education levels. The expected utility of attending a unilingual *L* school is then

$$U^{lu}(\mu) = -c + (p_l(1 - \mu_l^{-lu}) + p_{-l}(\mu_{-l}^b + \mu_{-l}^{lu}))(1 + \sigma) \quad \text{for } l = m, n. \quad (2)$$

The interpretation is that an individual pays c , becomes skilled, and speaks *Lish* when leaving school. She gets $1 + \sigma$ from production with the $p_l(1 - \mu_l^{-lu})$ native *Lish* speakers who have not attended a unilingual school in the other

is as in Lang (1986), where one group is exogenously assigned the role of the “economically dominant” group because its capital/labour ratio is assumed to be larger than that of the other group.

⁶Indeed, as shown by linguists (see e.g. Fishman 1977 for English and Hagège 1996 for France), one crucial factor behind language shift in populations over generations is the choice of the language(s) of instruction in school. In other words, languages not given the status of medium of instruction in primary school tend to be replaced by the language used in school. Here for simplicity we assume that this language shift takes place in the life span of one generation. Mass media, migrations and parental choices are other important factors behind language shift. For a dynamic setup in which the language spoken by the children (exogenously) depends on the language spoken by the parents and on the language spoken in the geographical location, see John and Yi (2001).

language and with the $p_{-l}(\mu_{-l}^b + \mu_{-l}^{lu})$ native speakers of the other language that have learnt *Lish* either in a bilingual school or in a unilingual *L* school.

Finally, an unskilled *L* has the same production partners as an individual attending a unilingual *Lish* school, as she speaks *Lish*. This individual saves on the cost of education but obtains a value of only 1 when producing:

$$\underline{U}_l(\mu) = p_l(1 - \mu_l^{-lu}) + p_{-l}(\mu_{-l}^b + \mu_{-l}^{lu}) \quad \text{for } l = m, n. \quad (3)$$

2.2 Education systems

An education system is defined as a menu of school type choices for each language group $L = \{M, N\}$. Although there are 49 possible education systems,⁷ Proposition 4 (see Section 5) shows that there is no system that the dominant prefer to the three following systems, to which we restrict our attention. Under *laissez-faire*, each individual taking education freely chooses whether to attend a unilingual *M* school, a unilingual *N* school, or a bilingual school. Under the *N bilingual* system, the *Ms* are restricted to unilingual *M* schools and the *Ns* to bilingual schools. Finally, under the *M unilingual* system, only unilingual *M* schools are allowed and so any individual attending school ends up speaking *Mish* only.

2.3 Equilibrium

The timing of the game is as follows. First, anticipating the future levels of education, the education system is chosen so as to maximise the expected utility of the dominant group. Second, each individual independently and simultaneously chooses whether to undertake education. Without loss of generality, we consider symmetric Nash equilibria in which all members of each group randomise between education and staying unskilled with the same probability.

3 Equilibrium education levels

3.1 Laissez-faire

If unilingual and bilingual schools cost the same, then any individual who invests in education will choose a bilingual school because it provides her with a second language and thus, *ceteris paribus*, enlarges her set of production partners. Then, under *laissez-faire* (denoted by *d* for “deregulated”), $\mu_l^{mu} = \mu_l^{nu} = 0$ and $\mu_l^d = \mu_l^b$ for $l \in \{m, n\}$. Subtracting (3) from (1) and using these expressions, the net incentive for taking education is here

$$\Delta U_l^d(\mu_{-l}^b) = -c + p_{-l}(1 - \mu_{-l}^b) + \sigma \quad \text{for } l = m, n; \quad (4)$$

that is, a native *Lish* speaker who undertakes education pays *c*, learns language $-L$, and thus gains as production partners the $p_{-l}(1 - \mu_{-l}^b)$ native $-L$ speakers

⁷There are seven possible arrangements for each group, namely: (i) unrestricted choice; access to any school (ii) except *N* unilingual, (iii) except *M* unilingual, or (iv) except bilingual; access only to (v) *M* unilingual, (vi) *N* unilingual, or (vii) bilingual. Seven arrangements for each of the two groups yields a total of 49 systems.

who do not speak *Lish* because they do not go to school. This is the communication effect of education. Moreover, there is a productivity gain σ from education that derives from the ability of the bilingual individual to produce an additional amount σ with any other individual.⁸

The Nash equilibria $(\gamma_n^d, \gamma_m^d) = (\gamma_n^b, \gamma_m^b)$ of this game are depicted in Figure 1. When the productivity gain covers the cost of education ($\sigma > c$), undertaking education is a dominating strategy for everybody. At the other extreme, if education is very expensive ($c > \sigma + \max\{p_m, p_n\}$), then any educational investment is a dominated strategy and nobody attends school.

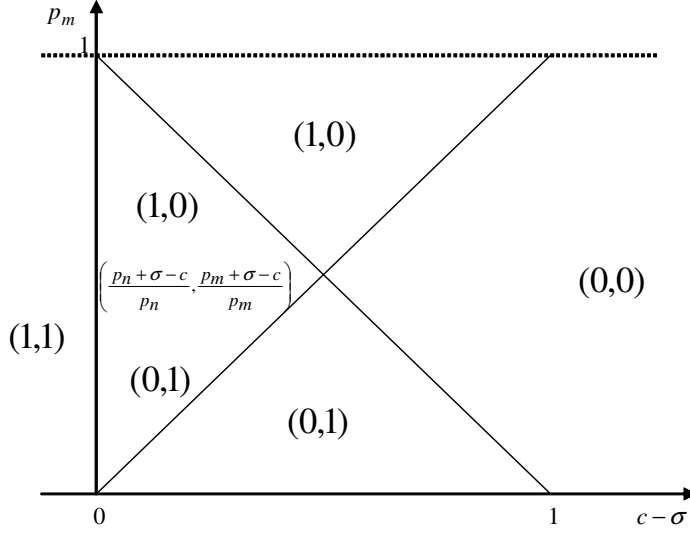


Figure 1 : Equilibrium education levels (γ_n^b, γ_m^b) under laissez-faire

For intermediate values of $c - \sigma$, intergroup interactions become relevant and the communication effect plays a role in equilibrium. From (4), the incentives to attend school for each group are decreasing in the education level of the other group, as learning the other language becomes less interesting for a higher prevalence of bilingualism in the other group (a “duplication” effect). If group $-L$ becomes fully educated, then all its members become bilingual, and thus can be reached by the uneducated L s. In this case, education simply affects the productivity of the L s ($\Delta U_l^d(1) = \sigma - c$ from (4)), and is unprofitable for them since $c > \sigma$. In turn, the absence of education among the L s makes group $-L$ willing to invest in schooling only if doing so generates a sufficient number p_l of additional production partners—that is, if $\Delta U_{-l}^d(0) = p_l + \sigma - c \geq 0$. Hence, $(1, 0)$ and $(0, 1)$ are equilibria if respectively $c - \sigma \in (0, p_m)$ and $c - \sigma \in (0, p_n)$. The two cases are not mutually exclusive, because both equilibria arise simultaneously if the two groups are roughly equal in size and education is not too expensive.⁹

⁸In countries where the language of the dominated has low prestige or where language groups have (strong) political identities, the dominant may not bother to learn carefully the language of the dominated when attending bilingual schools. In our model, if the dominant do not learn at all the language of the dominated then laissez-faire becomes identical to the bilingual system.

⁹ $\Delta U_l^d((p_{-l} + \sigma - c)/p_{-l}) = 0$ implies that there also exists an interior laissez-faire equilibrium, $((p_n + \sigma - c)/p_n, (p_m + \sigma - c)/p_m)$, in this case.

3.2 The N bilingual system

Under the N bilingual system (also referred to simply as the bilingual system), denoted by bi , the M s can attend only M unilingual schools and the N s can attend only bilingual schools.¹⁰ Thus, $\mu_m^b = \mu_m^{nu} = 0$, $\mu_n^{mu} = \mu_n^{nu} = 0$, $\mu_m^{bi} = \mu_m^{mu}$, and $\mu_n^{bi} = \mu_n^b$. Given these expressions, subtracting (3) from (1) for $l = n$ yields the net benefit from education for a native N ish speaker:

$$\Delta U_n^{bi} = -c + p_m + \sigma. \quad (5)$$

The individual pays c , reaches p_m additional partners as she learns M ish, and gains additional output σ , because she is now skilled and can produce with everybody.¹¹

In turn, subtracting (3) from (2) for $l = m$ shows that the net benefit from education for the M s is

$$\Delta U_m^{bi}(\mu_n^b) = -c + (p_m + p_n \mu_n^b) \sigma. \quad (6)$$

Because both the uneducated and educated M s speak M ish only and produce with the $p_m + p_n \mu_n^b$ M ish speakers, the impact of education for an M is here confined to the productivity effect σ .¹²

The Nash equilibria $(\gamma_n^{bi}, \gamma_m^{bi}) = (\gamma_n^b, \gamma_m^{um})$ of this game are depicted in Figure 2. When education is cheap ($c < \sigma$), the productivity effect alone is sufficient to render education a dominating strategy for the N s because by (5), $\Delta U_n^{bi} = -c + p_m + \sigma > 0$. Anticipating the high education levels of the N s, the skill effect is sufficiently strong to render education profitable also for the M s ($\Delta U_m^{bi}(1) > 0$). Hence, full education is the unique equilibrium in this case. As soon as the cost of education becomes larger than σ , education is a dominated strategy for the M s ($\Delta U_m^{bi}(\mu_n^b) < 0 \forall \mu_n^b$). Instead, the N s are still willing to pay for education if and only if the p_m new production possibilities it generates compensate for the insufficient productivity effect $c - \sigma$. Hence, the N s become bilingual provided $c - \sigma \in (0, p_m)$ and otherwise remain uneducated.¹³

¹⁰It would have been equivalent here to assume that the N s are free to choose between the three types of schools, since this would have implied that in equilibrium they choose bilingual schools. However, the two assumptions are no longer equivalent in Section 6.3, where we consider the case of higher cost of bilingual schools.

¹¹Comparing (5) with (4) for $l = n$, it appears that schooling incentives are higher here for the N s than under *laissez-faire*, since the M s never learn N ish and so schooling always enlarges the set of production partners for the N s. In contrast, under *laissez-faire* an uneducated N could still produce with the $p_m \mu_m^d$ bilingual M s.

¹²This implies that the representative M has lower incentives for education than under *laissez-faire* (see (4)), where education could also enlarge her set of production partners.

¹³Comparing Figure 2 with Figure 1, the N s take more education than under *laissez-faire* while the reverse is true for the M s.

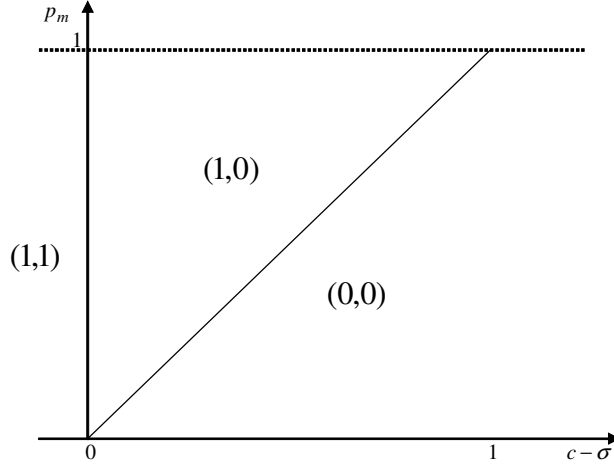


Figure 2 : Equilibrium education levels $(\gamma_n^b, \gamma_m^{\mu})$ under bilingualism

3.3 The M unilingual system

Under the M unilingual system (referred to simply as the unilingual system and denoted by *uni*), anyone undertaking education must go to a unilingual M school; in other words, $\mu_l^b = \mu_l^{nu} = 0$ and $\mu_l^{uni} = \mu_l^{mu}$ for $l = \{m, n\}$ by institutional design. Thus, just as under bilingualism, the M s never end up speaking Nish, and their net gain from education is still given by (6). Subtracting (3) from (2) for $l = n$, the net benefit of taking education for a representative N becomes

$$\Delta U_n^{uni}(\mu_n^{mu}) = -c + (p_m + p_n \mu_n^{mu})\sigma + p_m + p_n \mu_n^{mu} - p_n(1 - \mu_n^{mu}). \quad (7)$$

When attending school, this individual pays c , becomes skilled, and shifts language from Nish to Mish. In (7), the *productivity gain* from education is given by $(p_m + p_n \mu_n^{mu})\sigma$ —that is, the marginal value of education σ multiplied by the production partners (the M s and the other skilled N s) after schooling. In addition, education alters the set of production partners. This *communication effect* is captured by the remaining terms in (7). First, speaking Mish after school enables production with the p_m native Mish speakers and with the $p_n \mu_n^{mu}$ new Mish speakers. At the same time, the skilled N forgets Nish and thus can no longer produce with the $p_n(1 - \mu_n^{mu})$ unskilled N s.

Equation (7) generates an insight that is crucial to understanding the preferences over education systems. For the N s, attending school under unilingualism implies both becoming skilled and shifting language. Clearly, both features of unilingual schooling are more attractive the smaller the number of Nish speakers and, in particular, the larger the number of other N s attending the unilingual school. This positive communication externality is thus at the origin of a *bandwagon* or *snowball* effect in the schooling decisions of the N s. Indeed, it is easy to check from (7) that the net return to schooling for an N is increasing in the number of N s taking education (μ_n^{mu}). If the bandwagon effect is sufficiently strong, it generates multiple equilibria. In addition to the possibility of two extreme equilibria in which either all or none of the N s take education, an interior

equilibrium may exist.¹⁴ Intuitively, for the bandwagon effect to play a role in equilibrium, the dominated group must be sufficiently large; otherwise, avoiding school and restricting oneself to intragroup production is never profitable for the an N .

The Nash equilibria equilibria $(\gamma_n^{uni}, \gamma_m^{uni}) = (\gamma_n^{um}, \gamma_m^{um})$ of this game are depicted in Figure 3 (see Appendix A for full details). The $\bar{p}_m(c - \sigma)$ line characterises the critical size of the dominant group below which the bandwagon effect comes into play. It is upward sloping because education becomes less attractive as schooling costs increase, and thus remaining an N ish speaker is profitable even if the size of the N group shrinks. Above the $\bar{p}_m(c - \sigma)$ line, the dominant group is so large—and thus the productivity gain and communication effect are so strong relative to the cost of education—that education is a dominant strategy for the N s. Then, the M s become educated if schooling is sufficiently cheap ($c < \sigma$) and abstain from education otherwise.

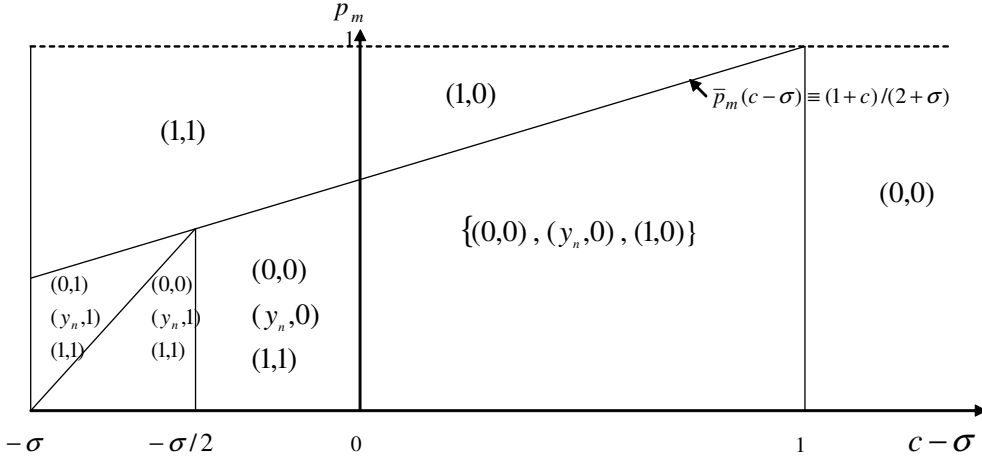


Figure 3: Equilibrium education levels $(\gamma_n^{mu}, \gamma_m^{mu})$ under the unilingual system, with $y_n \equiv (\bar{p}_m - p_m)/p_n$

Southeast of the $\bar{p}_m$ line in Figure 3, the bandwagon effect induces multiple equilibria. The fear of being caught alone as an N ish speaker induces high equilibrium education, but it may equally well be that an expected disinterest in education among the N s de facto discourages education.

4 Welfare

Expected welfare is obtained by adding up individual utility levels:

$$W(\mu) = \sum_{l=m,n} p_l [\mu_l^{mu} U^{mu}(\mu) + \mu_l^{nu} U^{nu}(\mu) + \mu_l^b U^b + (1 - \mu_l^{mu} - \mu_l^{nu} - \mu_l^b) \underline{U}_l(\mu)]. \quad (8)$$

In the presence of a benevolent social planner who is able to enforce welfare-maximising education levels under each system, the following proposition holds.

¹⁴This unstable equilibrium is sometimes referred to as a *tipping* equilibrium, a term coined by Schelling (1978).

Proposition 1 *Under centralisation, laissez-faire yields (weakly) higher expected welfare than any other education system.*

Proof. See Appendix B.1. ■

Given identical costs of unilingual and bilingual schools, the total expenditures associated with a given educational level are the same regardless of the education system. However, production (and thus expected utility) is larger the larger the proportion of bilingual individuals. Since bilingualism is maximised under laissez-faire, this system is always chosen by a central planner who can control education levels.

In reality, of course, no central planner can perfectly control the amount of effort that students spend on their studies—even in a system with mandatory education.¹⁵ In order to capture this degree of freedom, we consider a situation in which the central planner picks the educational system and then individuals decide whether or not to attend school. Under decentralised school attendance choice, the following proposition holds.

Proposition 2 *In the choice between laissez-faire, bilingualism, and unilingualism, the system yielding a higher decentralised welfare is depicted in Figure 4.*¹⁶

Proof. See Appendix B.2. ■

Figure 4 shows that laissez-faire ceases to be optimal in regions (I) and (IIa) owing to the different schooling incentives in the three systems. In region (I), the cost of education is high enough for it to be optimal that only the minority become educated: given $p_m > 0.5$, only the N s should attend school. However, under laissez-faire, the schooling incentives of the M s are strong if few of the N s are expected to take education. The majority (i.e., the wrong group) may thus end up undertaking education. In contrast, under the two other systems, the M s cannot learn N ish; this reduces their incentives to undertake education and guarantees that only the N s (if any) take education. In the choice between unilingualism and bilingualism, both yield full communication and exclusive minority education as the unique equilibrium in region (Ia). In region (Ib), bilingualism is preferred because the bandwagon effect (which comes into play when the dominated group is large) generates an additional no-education equilibrium in the unilingual system. In region (IIa), laissez-faire is suboptimal because it fails to generate sufficient schooling incentives. Yet the bandwagon effect can raise schooling under the unilingual system to a level that would be impossible to reach under either of the two other systems. Thus, even though bilingual schools are technologically superior, a system that restricts (some)

¹⁵There is, for example, a growing empirical literature on the effectiveness of financial incentives for school attendance in developing economies (see e.g. Bourguignon, Ferreira, and Leite 2003 and the references therein).

¹⁶The figure is drawn assuming $\sigma < 1$, although this assumption is irrelevant to the results we have obtained. Also, as previously shown, multiple equilibria sometimes arise under laissez-faire and under the unilingual system. In models with multiple equilibria, predictions generally depend on the equilibria under consideration. The ranking in this proposition builds on the exclusion of interior unstable equilibria in parts of the parameter range for which $p_m < 0.5$. All subsequent results hold for comparisons of all equilibria.

individuals from access to bilingual schools may be better because it provides stronger schooling incentives than laissez-faire.

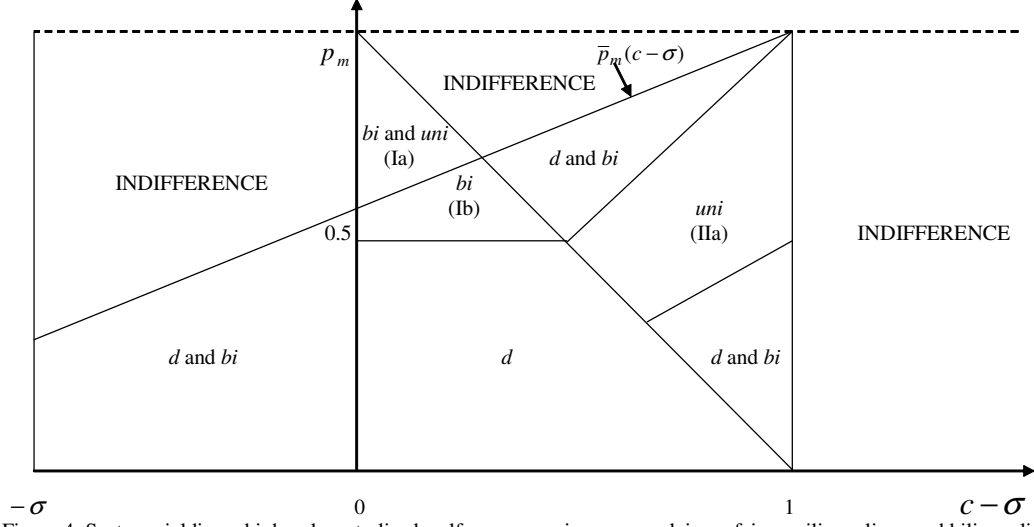


Figure 4: System yielding a higher decentralised welfare: comparison among laissez-faire, unilingualism, and bilingualism

Consider now *all* the possible education systems. An education system s is said to be *decentralised optimal* if there exists no alternative system that is (weakly or strictly) preferable to s in terms of decentralised welfare. Then, we can state the following proposition.

Proposition 3 *For $p_m > 0.5$, the education systems depicted in Figure 4 are decentralised optimal systems.*

Proof. See Appendix B.3. ■

When $p_m > 0.5$, the education systems depicted in Figure 4 attain the maximum centralised welfare level for $c < 1 + \sigma$ except in region (IIa). In that region, unilingualism generates the first-best welfare in one of its equilibria; however, since no system is able to reproduce the first-best level as a unique equilibrium, unilingualism is never outperformed. Finally, for $c > 1 + \sigma$, no system induces positive education and so all systems are equivalent.

The three systems may all be suboptimal whenever the dominant group is the minority. For example, the M bilingual system (where the M s become bilingual and the N s never learn M ish) generates the social optimum ($\gamma_n^{mu} + \gamma_n^{nu} + \gamma_n^b = 0, \gamma_m^b = 1$) as a unique equilibrium for $p_m \in (c - \sigma, 0.5)$ and $p_n > c - \sigma$ when neither of the three systems above is capable of doing so.

5 The choice of education system

This section examines how the M s and the N s rank different education systems, taking into account the equilibrium schooling levels under each system. In particular, we study whether language conflict can endogenously arise in our setup and whether this conflict is of the expected type—that is, a situation where the dominated favour the use of N ish in education while the dominant

oppose it. In addition, we determine which political rules, if any, enable society to reach a decentralised optimum.

The M s have been exogenously assigned the role of dominance and thus choose the system independently of whether they are a majority or a minority. Proposition 4 shows that there is no education system that the M s prefer to laissez-faire, the unilingual system, and the bilingual system, and determines when each of these three systems is chosen.

Proposition 4 *In decentralised equilibrium, the following statements hold. (i) If the cost of education is low ($c < \sigma$) or high ($c > 1 + \sigma$), then the dominant weakly prefer laissez-faire and bilingualism to any other education system, or are indifferent. (ii) For intermediate costs ($\sigma < c < 1 + \sigma$): (a) if the dominant group is sufficiently large ($p_m > c - \sigma$), then the dominant weakly prefer bilingualism to any other system or are indifferent; and (b) if the dominant group is sufficiently small ($p_m < c - \sigma$), then there exists no education system that the dominant prefer to the unilingual system.*

Proof. See Appendix B.4. ■

Every member of society would like to communicate with everybody else, since having fewer production partners for a given schooling investment can only reduce production opportunities and utility.

When education is cheap ($c < \sigma$), maximum communication is ensured under both laissez-faire and the N bilingual system because attending a bilingual school is a dominating strategy for the N s. When $c < \sigma$, education is profitable also for the M s, since the cost is always recovered via increased productivity. Instead, for instance, unilingualism fails to guarantee that all the N s attend school; hence the M s, who would not learn N ish, may not be able to produce with everybody and so do not choose that system.

When education becomes more expensive ($\sigma < c < 1 + \sigma$), the maximum gain from education σ becomes insufficient to cover the cost of education; then the M s try to eschew education while still producing with everybody. Laissez-faire is not always useful for this purpose when the dominant group is large ($p_m > c - \sigma$). Indeed, under this system, if the N s choose not to attend school then it is individually rational for each M to attend a bilingual school whenever the productivity gain coupled with the communication effect cover c . The dominant effectively commit to not undertaking education by choosing bilingualism, which legally prevents them from learning N ish. With this commitment device, intergroup production can be initiated only by the N s, which in turn lowers (raises) the education incentives of the M s (N s) and guarantees maximum communication paid for by the dominated. When the dominant group is small (i.e., $p_m < c - \sigma$), the bandwagon effect may generate the best outcome for the M s under unilingualism whereas laissez-faire and bilingualism fail to create sufficient incentives for the N s to undertake education.¹⁷ More generally, it can be shown that any other system has an equilibrium where the

¹⁷For $c - \sigma < 1 - p_m$ in this area, the choice of the M s between laissez-faire and unilingualism is indeterminate. Although the best outcome can be reached only under unilingualism, the N s prefer $(\gamma_n^d, \gamma_m^d) = (0, 1)$ to $(\gamma_n^{uni}, \gamma_m^{uni}) = (0, 0)$.

N s do not invest in education and so cannot be better than unilingualism from the M s' viewpoint.¹⁸

Proposition 5 characterises the preferences of the dominated with regard to education systems.

Proposition 5 *The dominated prefer laissez-faire to bilingualism (or are indifferent between the two) and prefer the bilingual to the unilingual system (or are indifferent between the two).*

Proof. See Appendix B.5. ■

Under laissez-faire, the M s learn N ish whenever they undertake education; thus the dominated can gain production partners without investing in education, which explains why they prefer laissez-faire to bilingualism. Instead, bilingualism and unilingualism have in common that the cost of intergroup communication is always borne by the N s. And among these two systems, the N s prefer bilingualism owing to the absence of negative network externalities and thus of bandwagon effects. In contrast, under unilingualism, the N s are in some cases “compelled” to undertake education even if this is relatively expensive, since doing so is the only way to learn M ish and to avoid remaining isolated from the rest of society.

The joint implication of Propositions 4 and 5 is that language conflict always comes with the desire of the M s to restrict the use of N ish in education and the opposition of the N s to this choice (see also Figure 5). This is an interesting result because language conflict is here of the expected type and does not rely on any direct utility that agents derive from speaking their own native language.¹⁹ In contrast, language conflict is an equilibrium phenomenon. The dominated want their language to be a means of instruction in school not because they are wedded to it but rather because abandoning it would force them to overinvest in education. Likewise, by legally restricting the amount of N speakers, the dominant maximise the incentive for learning M ish and thereby force the cost of communication on the dominated. In other words, the M s free-ride on the costs of speaking a common language, which are entirely borne by the N s.

More precisely, in region (I) of Figure 5, the dominant support a regulation that precludes them from learning N ish. Intuitively, the M s need a way of committing not to learn N ish in order for the N s to pay for education. This is not possible under laissez-faire, since it might be individually rational for the M s to become bilingual.²⁰ In contrast, the N s would like the M s to learn N ish so as to gain additional production partners without paying for education, and this is achieved under laissez-faire. In region (II), the dominant want to ban the use of N ish in education *for everybody*. As explained previously,

¹⁸Finally, if education is very costly ($c > 1 + \sigma$) then the choice of system does not matter because nobody ever goes to school.

¹⁹Adding an exogenous utility term of speaking one's mother tongue would only reinforce the language conflict. The dominated would have an additional reason for preferring bilingual schools, while the choice of the dominant would remain unaffected (since they maintain their mother tongue under all three systems).

²⁰In region (Ic), the M s prefer bilingualism to unilingualism in order to avoid the possible negative bandwagon effect under unilingualism. The bandwagon effect does not play in region (Ia), which explains why the M s are indifferent between the two systems in this region.

The graph illustrates the stability regions for different equilibria in the $(c - \sigma, p_m)$ plane. The horizontal axis represents $c - \sigma$ and the vertical axis represents p_m . The regions are defined by indifference curves and a vertical line at $c = 0$.

- (Ia) UNANIMITY d or bi** : Located in the top-left region, bounded by the p_m axis and the indifference curve $\bar{p}_m(c - \sigma)$.
- (Ic) UNANIMITY d or bi** : Located in the middle-left region, bounded by the p_m axis and the indifference curve $\bar{p}_m(c - \sigma)$.
- (II) UNANIMITY $Ms: uni$, $Ns: d$ or bi** : Located in the middle-right region, bounded by the indifference curve $\bar{p}_m(c - \sigma)$ and the vertical line at $c = 0$.
- (III) UNANIMITY d or bi** : Located in the bottom-left region, bounded by the p_m axis and the indifference curve $\bar{p}_m(c - \sigma)$.

The indifference curves are labeled **INDIFFERENCE** and **INDIFFERENCE $\bar{p}_m(c - \sigma)$** . The vertical line at $c = 0$ is labeled **INDIFFERENCE**.

Under a wide range of circumstances, there is no political tension over the choice of education system. When education is relatively cheap ($c < \sigma$) both groups either prefer laissez-faire or bilingualism or are indifferent between the three systems. Education is so cheap that it is important for both groups to generate the strongest possible incentive for schooling, thus avoiding potential negative bandwagon effects. For this reason, unilingualism is not chosen in region (III). When education is expensive ($c > 1 + \sigma$), nobody ever becomes educated and so the choice of system does not matter.

Proposition 6 *Under majority rule, an optimal decentralised system is always implemented. Under minority rule, the use of the dominated group language is too often restricted.*

15

6 Extensions

In this section, we extend the basic model to consider, in turn: (i) education subsidies; (ii) the possibility for native *Lish* speakers attending a $-L$ unilingual school to retain their mother tongue and thus become bilingual; (iii) an open economy with f_l *Lish* speakers abroad, $l = \{m, n\}$; (iv) an additional cost κ of attending bilingual schools; and (v) a combination of extensions (iii) and (iv). The expected utilities of attending the different types of school or remaining uneducated in extensions (ii)–(v) are presented in Appendix C; the full analytical details of all the extensions are contained in a Technical Appendix.²¹

6.1 Education subsidies

Failure to internalise communication externalities may lead to inefficient education decisions, as shown in Section 4. However, the inefficiencies can be overcome by an appropriate transfer system as follows.

Proposition 7 *There exists an education subsidy, targeted to the minority and financed by a proportional tax on production, that implements the socially optimal schooling level as the unique equilibrium under laissez-faire with decentralised schooling choice. However, the centralised optimum cannot always be implemented through a Pareto-improving policy.*

Proof. See the Technical Appendix. ■

The question of whether a transfer system can be voluntarily implemented under laissez-faire is related to whether or not it is Pareto-improving. As shown in the proof of Proposition 7, the proposed transfer system is strictly Pareto-improving whenever the source of the inefficiency is undereducation of both groups. In this case, we expect such a transfer system to be implemented under laissez-faire. If, instead, the problem lies in education of the wrong group (the majority), then there exists no combination of education subsidies and transfers (not even targeted lump-sum) that provide a Pareto improvement with respect to all (stable) laissez-faire equilibria. A Pareto improvement would require that the entire cost of the subsidy be borne by the minority. However, this is worse from the minority's viewpoint than an equilibrium in which the majority undertakes education. A Pareto-improving policy would therefore demand that the transfer system be made contingent on the laissez-faire equilibrium. This, in turn, would require an equilibrium refinement that selects among strict equilibria.

6.2 Mother tongue persistence

Imagine that the mother tongue is retained with probability α after unilingual education in the other language. In this case, every N attending a unilingual *Mish* school under unilingualism becomes bilingual with probability α and thus is subject to the bandwagon effect only with probability $1 - \alpha$.²² The result

²¹The Technical Appendix is available upon request and at <http://www.ifn.se/thomast>.

²²In equilibrium, nobody ever attends a unilingual school in the other language under laissez-faire or bilingualism, which implies that these two systems remain unchanged.

is a weaker bandwagon effect under unilingualism, which reduces the ability of this system to generate higher schooling levels than bilingualism or laissez-faire when education is expensive. However, unless persistence is perfect ($\alpha = 1$), the results remain qualitatively unchanged. The bandwagon effect survives, and there always exists an area in which the M s prefer unilingualism as well as an area in which unilingualism is a decentralised optimal system. Moreover, language conflicts follow the same pattern as in the benchmark case, and majority rule still implements a socially optimal system.

6.3 Cross-border spillovers

Consider an exogenous number f_l of Lish speakers abroad, $l = \{m, n\}$. Because the foreigners never shift language or become bilingual, their presence gives an additional advantage to bilingual schools (and thus to laissez-faire), since any system restricting access to bilingual schools in the home country eliminates production opportunities with foreigners. For instance, under unilingualism or bilingualism, the native M ish speakers can never produce with the f_n N ish-speaking foreigners. As a result, whenever the distribution abroad is skewed towards N ish, the following proposition holds.

Proposition 8 *If $f_n \geq \max\{p_m + f_m, (1 + f_m)/(2 + \sigma)\}$, both groups always (weakly) prefer laissez-faire to bilingualism and unilingualism.*

Proof. See the Technical Appendix. ■

The intuition is simple: For f_n sufficiently large, the M s choose the only system that allows them to learn N ish and thus reach the large number of N ish speaking foreigners, while the N s still have no incentive to restrict the use of their own language.

A second new result arises when f_n is large relative to p_n ($f_n > p_n/(1 + \sigma)$). In this case, any native N ish speaker retains many production partners abroad (f_n) even if the other native N ish speakers were to shift language by attending school under unilingualism, which implies that the bandwagon effect becomes irrelevant. Then unilingualism can no longer generate high schooling levels among the N s and ceases to be of interest to the dominant. The presence of a foreign language group may therefore serve to protect the language of a dominated minority.

In the rest of the cases (i.e., when the distribution abroad is balanced and f_n is small relative to p_n), most of the qualitative results of the baseline model still hold, as in this case opening the economy simply increases its size. In particular, unilingualism and bilingualism are chosen by the M s in some cases and can also be optimal in the absence of transfers, and the same type of language conflicts are observed. However, majority rule may result in the choice of a suboptimal system.²³

²³In particular, for $f_n > f_m$, the social planner may prefer the M s rather than the N s to invest in education (learn the other language) even if $p_m > 0.5$. Indeed, the higher educational costs associated with having $p_m > p_n$ individuals educated may be compensated for by a smaller loss of cross-border production opportunities (i.e., f_m times p_n instead of f_n times p_m). Meanwhile, the M s still prefer to free-ride on the education of the N s by choosing unilingualism or bilingualism.

6.4 Higher cost of bilingual schools

Whenever bilingual schools are more costly than unilingual schools ($\kappa > 0$), everybody taking education under laissez-faire chooses unilingual schools in stable equilibria. Indeed, since only corner equilibria are stable, any equilibrium with bilingual school attendance would be characterised by full communication. Then each individual who attends a bilingual school could save κ and keep the same production partners by shifting to a unilingual school in the language shared by everybody.

Although the central planner, given $\kappa > 0$, never chooses the bilingual system, the M s may still implement bilingualism because the extra cost of education falls on the N s. Excessive bilingualism occurs whenever the two groups are roughly equal in size and $p_m < 0.5$. Not only is the majority forced to learn the language of the minority in this case, but an additional distortion stems from this taking place in bilingual rather than unilingual schools. However, under majority rule, a suboptimal system is never chosen.

The sensitivity of bilingual school attendance under laissez-faire to the cost differential κ may cast doubt on the robustness of the baseline results. However, bilingual school attendance under laissez-faire is restored in an open economy whenever the differential cost κ of becoming bilingual is compensated for by a larger expansion in the set of production partners abroad. In this case, bilingual education may again be socially optimal, and the qualitative results of the benchmark model still hold. We thus view the closed economy with differential education cost as a case with too large a cost of learning an additional language.

7 Historical evidence

7.1 19th-century Belgium

Upon Belgium's independence in 1830, French had a predominant role even though French speakers were a minority.²⁴ In Flanders, Dutch was partly used as language of instruction in primary schools; however, secondary education was systematically provided in French until 1883, when a law established that some subjects in secondary public schools should be taught in Flemish. The initial predominance of French has been explained by the fact that "[the] bourgeoisie was overwhelmingly French-speaking, even in the Flemish provinces" (McRae 1986, p. 21) and that only 46,000 individuals in a population of about four million were given the right to vote. In our model, this can be interpreted as a dominant minority's choice of unilingualism.

7.2 France

Language policy was an important political issue during the French Revolution (1789–1794), and a series of French unilingual decrees were approved in 1794 by the radical revolutionaries (so-called *montagnards*; Hagège 1996). Although these decrees did not survive the fall of the *montagnards*, they became the foundations of French language policy (Weber 1976). In 1794, only about 40%

²⁴Dutch (Flemish) speakers accounted for 57% of the population in 1846 (McRae 1986).

Table 1: Proportion of French unilingual public schools at the *département* level in France (1863)

Dependent variable: Proportion of French unilingual public schools			
French speakers in the population	1.17** (.442)	0.805 (.679)	0.210 (.753)
Amount paid for education by a family	0.318*** (.108)	0.291** (.117)	0.267** (.117)
Log income per head	0.960* (.482)		1.2 (1.24)
Unilingual French-speaking <i>département</i> (dummy)		3.03*** (.584)	2.97*** (.795)

Notes: Figures reported are the coefficients obtained from tobit estimation. Standard errors are in parentheses, and *, **, and *** denote significance at 10%, 5%, and 1% levels, respectively. Data on language are from the *Archives Nationales* and can be found in Weber (1976). Data on average income levels are from Vapereau (1867) and the remaining data are from Ministère de l’Instruction Publique (1878). All data are for 1863.

of the population were native French speakers (Calvet 2002, p. 218).²⁵ Among the other language groups, the largest was Occitan and next came Breton and Alsatian.²⁶

Our model predicts that minority size may not be the most relevant variable for understanding the choice of education system. Indeed, inspecting Figure 5 reveals that, in the absence of information concerning the net cost $c - \sigma$ of education, the size p_m of the dominant group does not determine whether unilingual or bilingual schools are chosen at equilibrium. Using data from Weber (1976) for 1863, we can compute the proportion of public schools using only French in each of the 89 *départements* as well as the proportion of French speakers in the local population. The data show that there was regional variation in educational systems before the introduction of the 1880–1882 Ferry Laws, which instituted free primary education and legally established French as the only language of instruction in schools (Chervel 1992).

We next regress the proportion of French unilingual schools on a number of *département*-level variables for the 89 *départements*. The results are reported in Table 1. Column 1 shows that there is a positive relationship between the proportion of French speakers in the population and the proportion of French unilingual schools. In addition, the proportion of French unilingual schools is positively related to the average direct cost of education for parents in each *département*. This may indicate that parents were willing to invest more in education if schools were in French—most likely by a social mobility argument, since a knowledge of French was necessary in skilled occupations.

However, this first regression does not take into account that, in the 55 fully French-speaking *départements*, the possibility of having non-French-speaking schools was not even considered. We control for this by introducing a dummy variable for unilingual French-speaking *départements* in the regressions of columns

²⁵This estimate is based on the language report conducted by Grégoire (1794).

²⁶Additionally, small minorities were speaking Franco-provençal, Basque, Catalan, Corsican, or Flemish. Each *département* (with the exception of the Basses-Pyrénées) had at most two language groups.

2 and 3. The results so obtained show that the relationship between the proportion of French speakers in the population and the proportion of French unilingual schools is no longer significant.

7.3 Finland

The current institutional language framework in Finland was established by the Constitution of 1919 and a series of language laws, the most important of which was approved in 1922. In 1920, Swedish speakers constituted only 11% of the population (McRae 1997); the rest of the population spoke Finnish, except for a tiny Sami group. Nevertheless, the Constitution recognised Finnish and Swedish as national languages on an equal basis. The language clauses of the Constitution were approved by overwhelming majorities that ranged from 88% to 96% (Eduskunta-Riksdag 1920, pp. 1028–30), with support coming from both language groups. The few deputies that opposed the approved system belonged to Finnish-speaking parties and supported a system closer to Finnish unilingualism (see Jackson 1938 and McRae 1997). At the same time, the parties with support from the Swedish speakers never sustained Swedish unilingualism, supporting instead the proposed symmetric bilingual system. For this reason, we argue that the Finnish speakers at that time were (in terms of our model) the dominant group, despite the historical overrepresentation of Swedish speakers among the elite.

The educational system is such that each municipality must provide schooling in the minority language (Swedish or Finnish) once a certain number of parents require it. Given the symmetry of the system, the unanimity for bilingual schools in Finland may be viewed in terms of our model as unanimity for *laissez-faire*, and it can be explained in an economy with cross-border spillovers in light of Proposition 8. According to this proposition, if the language distribution abroad is skewed towards the language of the dominated (here, the Swedish speakers), then *laissez-faire* is always unanimously chosen. Taking into account the number of Finnish and Swedish speakers in Finland and Sweden in 1920, we have $p_m = 0.89$, $p_n = 0.11$, $f_m = 0.0097$ and $f_n = 1.89$; hence the condition in Proposition 8 on the skewness of the language distribution is satisfied even for $\sigma = 0$.²⁷

8 Conclusion and further discussion

Although many countries are multilingual or have been historically formed by several language groups, language diversity has not always lead to language conflict between the groups when deciding on the language of instruction in school; nor has it always been the case that both language groups have agreed upon a unilingual or a bilingual system. One way to understand this variety of outcomes is to assume that agents enjoy some utility from speaking their own language and that compromise over language issues through political bargaining is reached only in some cases.

²⁷In 1920, Finland was inhabited by 341,000 Swedish speakers and 2,764,000 Finnish speakers (McRae 1997, p. 86); Sweden had 5,904,489 inhabitants (Statistiska Centralbyrån 2006), of which 30,247 were Finnish-speakers (Statistiska Centralbyrån 1924, p. 14).

In this paper, we take a different stand on the issue and propose a model centered on individual incentives to attend school. We show that these incentives vary across groups depending on the nature of the education system. In particular, education systems restricting the use of the language of the politically dominated group produce stronger incentives for the dominated to attend school. As a result, the dominated either bear a large part of the costs associated with adopting a common language or actively defend the use of their mother tongue in order to avoid paying such costs. In contrast, the dominant may defend such a restriction in order to free-ride on the educational investment of the dominated. Thus language conflict of the expected type is shown to arise endogenously as the result of economic conflict.

Should we expect language conflicts to result in the adoption of a suboptimal language system? According to our model, the answer crucially depends on the nature of political institutions. Specifically, we show that democratic institutions (interpreted as majority rule) choose an optimal education system, whereas autocracies too often restrict the use of the dominated group’s language.

Our model delivers a number of empirically grounded results via a parsimonious set of economic and political features, but there are aspects of existing language policies that can be rationalised only by further modeling assumptions. In particular, our model does not address the within-country regional distribution of languages and does not allow for a two-tier political decision process—say at the local and federal level. Some countries admit multiple languages of instruction while restricting access to specific languages on a geographical basis. According to this territoriality principle, the law specifies the linguistic boundaries inside the country and provides each territory with instruments for retaining its legally defined language(s). In Switzerland, the territoriality principle was established by its 1848 Constitution in order to neutralise possible effects of the newly declared freedom of movement about the country on the language composition of cantons (McRae 1983). At the same time, the Constitution instituted the equality of German, French, and Italian together with the cantonal choice of languages. Based on this institutional equilibrium which is still in force today (except for a greater role given to Romansh), some cantons with sizeable language minorities have implemented unilingual education systems. In Belgium, a similar system was instituted by the 1962–63 language laws: currently Dutch is the unique language of instruction in Flanders and the same applies for French in Wallonia, while French and Dutch coexist as languages of instruction in the Brussels capital region.²⁸ A framework that accounts for such institutional choices should include geographical distance and the cross-regional migration decisions of individuals.

Language choice in education can also lead to conflicts between different political decision levels, as in the case of Canada. In 1977, Bill 101 in Quebec established that “only children whose father or mother received most of their primary education in English in Quebec have access to English schools” (Barbaud 1998, p. 185). However, this was overturned by a new active bilingual

²⁸There are exceptions in some small specific territories of Flanders and Wallonia, where both French and Dutch are used as languages of instruction up to primary school. In the German-speaking cantons, German and French are the languages of instruction.

policy at the federal level (Canada Constitution Act of 1982) which established that all Canadian citizens whose mother tongue is French or English—or who have received their primary education in Canada in one of these two languages—have the right to have all their children educated in that same language (when the number of children so warrants). The potential conflict between federal and regional institutions in the determination of language policies poses an interesting direction for future research.

Appendix A. Equilibrium education levels under unilingualism

A.1. Education levels of the N s

Define $\bar{p}_m \equiv (1 + c)/(2 + \sigma)$; then substitute into (7) and rewrite to get $\Delta U_n^{uni}(\mu_n^{um}) = (2 + \sigma)(p_n \mu_n^{um} + p_m - \bar{p}_m)$. Now $c > 1 + \sigma$ implies $\bar{p}_m > 1$, so $\Delta U_n^{uni}(\mu_n^{um}) < 0$ and hence $\gamma_n^{uni} = 0$ in this case. Next, $c < 1 + \sigma$ implies $\bar{p}_m < 1$ and so education is a strictly dominating strategy for all $p_m > \bar{p}_m$; hence $\gamma_n^{uni} = 1$ in this case. If $c < 1 + \sigma$ and $p_m < \bar{p}_m$, then $\Delta U_n^{uni}(1) = (2 + \sigma)(1 - \bar{p}_m) > 0$, $\Delta U_n^{uni}(0) = (2 + \sigma)(p_m - \bar{p}_m) < 0$, and $\Delta U_n^{uni}((\bar{p}_m - p_m)/p_n) = 0$ together imply that all $\gamma_n^{uni} \in \{0, (\bar{p}_m - p_m)/p_n, 1\}$ are equilibria.

A.2. Education levels of the M s

By (6), $\Delta U_m^{uni}(\mu_n^{um}) \leq \Delta U_m^{uni}(1) = \sigma - c$ implies that $\gamma_m^{uni} = 0$ is the unique equilibrium for $c > \sigma$. Next, $\Delta U_m^{uni}(\mu_n^{um}) \geq \Delta U_m^{uni}(0) = p_m \sigma - c$ implies $\gamma_m^{uni} = 1$ for all $p_m > c/\sigma$. For $c < \sigma$ and $p_m < c/\sigma$, the equilibrium education level γ_m^{uni} depends on γ_n^{uni} . First, $\gamma_n^{uni} = 1$ implies $\gamma_m^{uni} = 1$ since $\Delta U_m^{uni}(1) = \sigma - c > 0$. Second, $\gamma_n^{uni} = 0$ implies $\gamma_m^{uni} = 0$ since $\Delta U_m^{uni}(0) = (p_m - c/\sigma)\sigma < 0$. Finally, $\Delta U_m^{uni}((\bar{p}_m - p_m)/p_n) = \bar{p}_m \sigma - c = (\sigma - 2c)/(2 + \sigma)$. Therefore, $\gamma_n^{uni} = (\bar{p}_m - p_m)/p_n$ implies that $\gamma_m^{uni} = 1$ for $c < \sigma/2$ and that $\gamma_m^{uni} = 0$ for $c > \sigma/2$.

Appendix B. Proofs

Let Γ^s be the set of equilibria under education system s and let $\gamma^s \in \Gamma^s$ be a specific equilibrium under s . Let $u_l^s(\gamma^s) = \gamma_l^{mu} U^{mu}(\gamma^s) + \gamma_l^{nu} U^{nu}(\gamma^s) + \gamma_l^b U^b + (1 - \gamma_l^b - \gamma_l^{mu} - \gamma_l^{nu}) \underline{U}_l(\gamma^s)$ be the indirect utility for a native *Lish* speaker under s given the choice of γ^s . Group L is said to *weakly prefer* system s to system s' (denoted $s \succeq_l s'$) if at least one of them displays multiple equilibria and if $\min\{u_l^s(\gamma^s) \mid \gamma^s \in \Gamma^s\} \geq \max\{u_l^{s'}(\gamma^{s'}) \mid \gamma^{s'} \in \Gamma^{s'}\}$. The preference is strict ($s \succ_l s'$) if the inequality is strict—no matter how many equilibria the two systems generate. Finally, a native *Lish* speaker is indifferent between the two systems ($s \sim_l s'$) if all equilibria generate $u_l^s(\gamma^s) = u_l^{s'}(\gamma^{s'})$. Subscript w on the preference ordering denotes the social welfare ordering; $W(\gamma^s)$ denotes expected welfare, as defined in (8) for equilibrium γ^s . If $s \in \{d, uni, bi\}$ then expected welfare is denoted $W^s(\gamma_n^s, \gamma_m^s)$, with $(\gamma_n^d, \gamma_m^d) = (\gamma_n^b, \gamma_m^b)$, $(\gamma_n^{uni}, \gamma_m^{uni}) =$

$(\gamma_n^{mu}, \gamma_m^{mu})$, and $(\gamma_n^{bi}, \gamma_m^{bi}) = (\gamma_n^b, \gamma_m^{mu})$. In addition, $\mathbf{x}$ denotes the optimal (centralised) education levels.

B.1. Proof of Proposition 1

For any distribution $\boldsymbol{\mu}$ of education levels, we show that there exists an alternative distribution $\tilde{\boldsymbol{\mu}}$ under d such that $W(\tilde{\boldsymbol{\mu}}) \geq W(\boldsymbol{\mu})$. Pick any $\boldsymbol{\mu}$ and consider $\tilde{\boldsymbol{\mu}}$, where a fraction $\tilde{\mu}_l = \mu_l = \mu_l^b + \mu_l^{mu} + \mu_l^{nu}$ of the L s ($l \in \{m, n\}$) are forced to take education but are free to choose the type of school. Naturally, everybody chooses a bilingual school under d . Hence, $\tilde{\mu}_l^{mu} = \tilde{\mu}_l^{nu} = 0$ and $\tilde{\mu}_l^b = \tilde{\mu}_l = \mu_l$ for all $l \in \{m, n\}$. The welfare difference is

$$\begin{aligned} W(\tilde{\boldsymbol{\mu}}) - W(\boldsymbol{\mu}) &= \sum_{l=m,n} (p_l \mu_l^{lu} + p_{-l} \mu_{-l}^{lu}) (U^b - U^{lu}(\boldsymbol{\mu})) \\ &\quad + \sum_{l=m,n} p_l (1 - \mu_l) (\underline{U}_l(\tilde{\boldsymbol{\mu}}) - \underline{U}_l(\boldsymbol{\mu})). \end{aligned}$$

Now $U^b \geq U^{lu}(\boldsymbol{\mu})$ and $\underline{U}_l(\tilde{\boldsymbol{\mu}}) - \underline{U}_l(\boldsymbol{\mu}) = p_m \mu_m^{-lu} + p_n \mu_n^{-lu} \geq 0$ for $l \in \{m, n\}$ imply that $W(\tilde{\boldsymbol{\mu}}) \geq W(\boldsymbol{\mu})$. We determine $\mathbf{x} = (x_n^d, x_m^d)$ by plugging $\mu_l^{mu} = \mu_l^{nu} = 0$ and $\mu_l^b = \mu_l^d$ for $l \in \{m, n\}$ into (8) to derive the laissez-faire welfare $W^d(\mu_n^d, \mu_m^d) = 1 - 2p_m p_n (1 - \mu_m^d)(1 - \mu_n^d) + (\sigma - c)(p_m \mu_m^d + p_n \mu_n^d)$. Then $W^d(1, 1) - W^d(\mu_n^d, \mu_m^d) = 2p_m p_n (1 - \mu_m^d)(1 - \mu_n^d) + (\sigma - c)(p_m(1 - \mu_m^d) + p_n(1 - \mu_n^d))$ implies that $\mathbf{x} = (1, 1)$ for all $c \leq \sigma$. Similarly, $W^d(1, 0) - W^d(\mu_n^d, \mu_m^d) = \mu_m^d(c - \sigma)(2p_m - 1 + p_n \mu_n^d) + p_n(2p_m + \sigma - c)(1 - \mu_m^d)(1 - \mu_n^d)$ implies that $\mathbf{x} = (1, 0)$ for all $c - \sigma \in (0, 2p_m)$ and $p_m > 0.5$. The proof that $\mathbf{x} = (0, 1)$ if $c - \sigma \in (0, 2p_n)$ and $p_n > 0.5$ is analogous. Finally, $W^d(0, 0) - W^d(\mu_n^d, \mu_m^d) = p_m \mu_m^d(c - \sigma - 2p_n) + p_n \mu_n^d(c - \sigma - 2p_m) + 2p_m p_n \mu_m^d \mu_n^d$ implies that $\mathbf{x} = (0, 0)$ for all $c - \sigma > 2 \max\{p_m, p_n\}$.

B.2. Proof of Proposition 2

The equilibrium education levels $\gamma^s \in \Gamma^s$ for $s \in \{d, uni, bi\}$ are taken from Section 3 and $\mathbf{x}$ from the preceding proof.

(i) For $c < \sigma$, d and bi both uniquely implement maximal communication and education; so does uni , provided also that $p_m > \bar{p}_m \equiv (1 + c)/(2 + \sigma)$. Hence, $W^d = W^{bi} = W^{uni} = W^d(\mathbf{x})$ in this case. If $p_m < \bar{p}_m$ then uni also generates undereducation equilibria and so ceases to be optimal.

(ii) For $c > 1 + \sigma$, no system generates positive education and so all systems are equivalent.

(iiia) Let $\sigma < c < 1 + \sigma$ and $p_m > 0.5$. If $p_m > \bar{p}_m$ and $p_n < c - \sigma$ also hold, then all three systems uniquely implement maximal communication through maximal (no) education of the minority (majority) in the majority language—that is, all systems attain $W^d(\mathbf{x})$. If instead $p_m > \bar{p}_m$ but $p_n > c - \sigma$ (region Ia), then d may educate the wrong group and thus is no longer optimal. For $p_n < c - \sigma$ and $p_m \in (c - \sigma, \bar{p}_m)$, uni may lead to undereducation and hence is no longer optimal. Fourth, if $p_m \in (1/2, \bar{p}_m)$ and $p_n > c - \sigma$ (region Ib), then of all the three systems only bi uniquely attains $W^d(\mathbf{x})$. Finally, if $p_m \in (1/2, c - \sigma)$ (a subset of region II), then we have $\Gamma^{bi} = \Gamma^d = (0, 0)$ and $\Gamma^{uni} = \{(1, 0), ((\bar{p}_m - p_m)/p_n, 0), (0, 0)\}$. Given that $W^{uni}(1, 0) = W^d(\mathbf{x})$,

$W^{uni}((\bar{p}_m - p_m)/p_n, 0) - W^{uni}(0, 0) = (p_m - p_n)(\bar{p}_m - p_m) > 0$, and $W^{uni}(0, 0) = W^{bi}(0, 0) = W^d(0, 0)$, it follows that $uni \succeq_w bi \sim_w d$.

(iiib) If $\sigma < c < 1 + \sigma$ and $p_m < 0.5$, then d is the only system of the three that can possibly implement $\mathbf{x}$ as an equilibrium; this happens for $p_n > c - \sigma$. Moreover, if $p_m < c - \sigma$ then d uniquely implements $\mathbf{x}$, but if $p_m > c - \sigma$ then d generates multiple equilibria. Restricting attention to stable equilibria, we can still rank the three systems. $W^d(1, 0) = W^{uni}(1, 0) = W^{bi}(1, 0)$ and $W^s(1, 0) - W^s(0, 0) = p_n(2p_m + \sigma - c) > 0$ for $s = uni, bi$ implies $d \succeq_w uni$ and $d \succeq_w bi$. Next, $\Gamma^{bi} = \Gamma^d = \{0, 0\}$ if $p_n < c - \sigma$. Given that $W^{bi}(0, 0) > W^{uni}(1, 0)$ if and only if $p_m < (c - \sigma)/2$, it follows that $d \sim_w bi \succeq_w uni$ for $p_m < (c - \sigma)/2$ and that $uni \succeq_w bi \sim_w d$ for $p_m > (c - \sigma)/2$ if we are considering stable equilibria only (we are then back to region II).

B.3. Proof of Proposition 3

The preceding proof shows that the systems depicted in Figure 4 all attain $W^d(\mathbf{x})$ through a unique equilibrium whenever $p_m > 0.5$ and $c < 1 + \sigma$, except for $p_m \in (1/2, c - \sigma)$. As a result, they cannot be outperformed by any other system. For $p_m \in (1/2, c - \sigma)$ we have $W^{uni}(1, 0) = W^d(\mathbf{x})$, implying that any system that outperforms uni must uniquely implement $W^d(\mathbf{x})$. However, as long as $p_m < c - \sigma$, any system will necessarily have $\tilde{\gamma}_n = \tilde{\gamma}_n^b + \tilde{\gamma}_n^{mu} + \tilde{\gamma}_n^{nu} = 0$ as part of a stable equilibrium. Indeed, $\Delta U_n^{lu}(\tilde{\gamma}) \leq \Delta U_n^b(\tilde{\gamma})$ always holds (for $l = m, n$) and $\Delta U_n^b(\tilde{\gamma}) = \sigma + p_m - c - p_m(\tilde{\gamma}_m^b + \tilde{\gamma}_m^{nu}) < 0$ for $p_m < c - \sigma$ and $\tilde{\gamma}_n = 0$. Therefore, uni is a decentralised optimum in this case. Finally, no system generates positive education for $c > 1 + \sigma$; hence, in this case, all systems are equivalent in welfare.

B.4. Proof of Proposition 4

Compare an equilibrium γ^s under system $s \in \{d, uni, bi\}$ such that $\gamma_n^{mu} + \gamma_n^b = 1$ to an arbitrary equilibrium $\tilde{\gamma}$ under system s' , where γ_l^s and $\tilde{\gamma}_l$ denote the aggregate education of group L in each of the equilibria. Observe first that $U_m^{mu}(\gamma^s) = U^b = 1 + \sigma - c$ and that $U_m(\gamma^s) = 1$. Some algebraic manipulations yield $u_m^s(\gamma^s) - u_m^{s'}(\tilde{\gamma}) = (\gamma_m^s - \tilde{\gamma}_m)\Delta U_m^s(1) + \tilde{\gamma}_m^{nu}(p_m(1 - \tilde{\gamma}_m^b - \tilde{\gamma}_m^{nu}) + p_n\tilde{\gamma}_n^{mu})(1 + \sigma) + (\tilde{\gamma}_m^{nu}(1 + \sigma) + 1 - \tilde{\gamma}_m)(p_m\tilde{\gamma}_m^{nu} + p_n(1 - \tilde{\gamma}_n^b - \tilde{\gamma}_n^{mu}))$, which is non negative because $\gamma_m^s < \tilde{\gamma}_m \leq 1$ implies $\Delta U_m^s(1) \leq 0$ and $\gamma_m^s > \gamma_m \geq 0$ implies $\Delta U_m^s(1) \geq 0$. Hence γ^s is the upper bound to M 's equilibrium utility.

Part (i): For $c < \sigma$, $\gamma_n^b = 1$ is reached as a unique equilibrium both under bi and under d ; consequently, the two systems are equivalent. For $c > 1 + \sigma$, all systems are characterised by zero education and thus are equivalent.

Part (ii), $\sigma < c < 1 + \sigma$: For $p_m > c - \sigma$, $\gamma_n^b = 1$ is reached as a unique equilibrium under bi . For $p_m < c - \sigma$, $\gamma_n^{mu} = 1$ and zero education in both groups are the two stable unilingual equilibria. In order for a competing system to outperform uni , it must have $\tilde{\gamma}_n^{mu} + \tilde{\gamma}_n^b = 1$ in every equilibrium for $p_m < c - \sigma$. However, any education system necessarily has $\tilde{\gamma}_n = 0$ as a (stable) equilibrium in this interval (see the preceding proof).

B.5. Proof of Proposition 5

Consider the arbitrary equilibria γ^d , γ^{uni} , and γ^{bi} . Manipulating terms yields $u_n^d(\gamma^d) - u_n^{bi}(\gamma^{bi}) = (\gamma_n^d - \gamma_n^{bi})\Delta U_n^d(\gamma_m^d) + p_m\gamma_m^d(1 - \gamma_n^{bi}) \geq 0$ and $u_n^{bi}(\gamma^{bi}) - u_n^{uni}(\gamma^{uni}) = (\gamma_n^{bi} - \gamma_n^{uni})\Delta U_n^{bi}(\gamma^{bi}) + p_n\gamma_n^{uni}(1 - \gamma_n^{uni})(2 + \sigma) \geq 0$, which imply that $u_n^d(\gamma^d) \geq u_n^{bi}(\gamma^{bi}) \geq u_n^{uni}(\gamma^{uni})$ for all γ^d, γ^{uni} , and γ^{bi} .

Appendix C. Extensions

For a native Lish speaker, the expected utility from attending a bilingual school, a unilingual Lish school, and a unilingual school in the other language ($-L$) are given as follows:

$$\begin{aligned} U^b &= -c - \kappa + (1 + f_m + f_n)(1 + \sigma); \\ U_l^{lu}(\mu) &= -c + [p_l(1 - (1 - \alpha)\mu_l^{-lu}) + p_{-l}(\mu_{-l}^b + \mu_{-l}^{lu}) + f_l](1 + \sigma); \\ U_l^{-lu}(\mu) &= -c + \alpha(1 + f_m + f_n)(1 + \sigma) \\ &\quad + (1 - \alpha)[p_l(\mu_l^b + \mu_l^{-lu}) + p_{-l}(1 - (1 - \alpha)\mu_{-l}^{lu}) + f_{-l}](1 + \sigma). \end{aligned}$$

For this individual, the utility of *not* taking education is

$$\underline{U}_l(\mu) = p_l(1 - (1 - \alpha)\mu_l^{-lu}) + p_{-l}(\mu_{-l}^b + \mu_{-l}^{lu}) + f_l.$$

References

- Barbaud, Philippe (1998). “French in Quebec.” In *Language in Canada*, edited by John Edwards. Cambridge University Press.
- Bourguignon, François, Francisco Ferreira, and Philippe Leite (2003). “Conditional Cash Transfers, Schooling, and Child Labor: Micro-Simulating Bolsa Escola.” *The World Bank Economic Review*, 17, 229-254.
- Calvet, Louis-Jean (2002). *Linguistique et colonialisme*. Editions Payot et Rivages.
- Card, David (1999). “The Causal Effect of Education on Earnings.” In *The Handbook of Labor Economics*, vol 3A, edited by Orley Ashenfelter and David Card. North Holland.
- Chervel, André (1992). *L’enseignement du français à l’école primaire. Textes officiels*, Tome 1, 1791-1879. Economica.
- Chiswick, Barry R., and Paul W. Miller (1995). “The Endogeneity between Language and Earnings: International Analyses.” *Journal of Labor Economics*, 13, 246-88.
- Church, Jeffrey, and Ian King (1993). “Bilingualism and network externalities.” *Canadian Journal of Economics-Revue canadienne d’Economie*, 25, 337-345.
- Eduskunta-Riksdag (1920). Valtiopäivät 1919. Pöytäkirhat I. Istunnot 1-53. Helsinki.

- Encyclopaedia Britannica (2003). *Britannica Book of the Year 2002*.
- Farrell, Joe, and Paul Klemperer (2007). "Coordination and Lock-In: Competition with Switching Costs and Network Effects." In *Handbook of Industrial Organization*, Vol 3, edited by Mark Armstrong and Robert Porter. North-Holland.
- Fishman, Joshua (1977). "The Spread of English as a New Perspective for the Study of Language Maintenance and Language Shift." In *The spread of English: the Sociology of English as an Additional Language*, edited by Joshua Fishman, Robert Cooper and Andrew Conrad. Rowley, Mass.: Newbury House.
- Ginsburgh, Victor, Ignacio Ortuno-Ortín, and Shlomo Weber (2005). "Disenfranchisement In Linguistically Diverse Societies: The Case of The European Union." *Journal of the European Economic Association*, 3, 946-965.
- Grégoire, Henri (1794). "Sur la nécessité et les moyens d'annéantir les patois et d'universaliser l'usage de la langue française." In *Une politique de la langue. La Révolution Française et les patois: l'enquête Grégoire*, edited by Michel de Certeau, Dominique Julia and Jacques Revel. Folio, Gallimard. 1975.
- Hagège, Claude (1996). *Le Français, histoire d'un combat*. Le livre de poche.
- Jackson, John H. (1938). *Finland*, George Allen and Unwin Ltd.
- John, Andrew, and Kei-Mu Yi (2001). "Language and Location." Mimeo. INSEAD, Paris.
- Laitin, David D. (1994). "The Tower of Babel as a Coordination Game: Political Linguistics in Ghana." *American Political Science Review*, 88, 622-634.
- Lang, Kevin (1986). "A Language Theory of Discrimination." *Quarterly Journal of Economics*, 101, 363-382.
- Lazear, Edward P. (1999). "Culture and Language." *Journal of Political Economy*, 107, 95-126.
- McRae, Kenneth D. (1983). *Conflict and Compromise in Multilingual Societies. Vol. 1, Switzerland*. Wilfried Laurier University Press.
- McRae, Kenneth D. (1986). *Conflict and Compromise in Multilingual Societies. Vol. 2, Belgium*. Wilfried Laurier University Press.
- McRae, Kenneth D. (1997). *Conflict and Compromise in Multilingual Societies. Vol. 3, Finland*. Wilfrid Laurier University Press.
- Méltiz, Jacques (2002). "Language and Foreign Trade." *CEPR Working Papers* No. 3590 (October). London.
- Ministère de l'Instruction Publique (1878). *Statistique de l'enseignement primaire. Tome second, Statistique comparée de l'enseignement primaire, 1829-1877*. Paris: Imprimerie Nationale.

Pool, Jonathan (1991). "The Official Language Problem." *American Political Science Review*, 85, 495-514.

Schelling, Thomas C. (1978). *Micromotives and Macrobehavior*. New York: W.W. Norton.

Statistiska Centralbyrån (1924). *Statistisk årsbok för Sverige*. Stockholm.

Statistiska Centralbyrån (2006). <http://www.ssd.scb.se>

Vapereau, G. (1867). *Atlas géographique, statistique et historique des départements de la France*. Paris: Bureau du Magasin du Foyer.

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