Unilingual versus bilingual education: a political economy analysis*

Javier Ortega
Université de Toulouse (GREMAQ, IDEI),
CEP (LSE), and CEPR, IZA

Thomas P. Tangerås
Research Institute of Industrial Economics (IfN), Stockholm

February 27, 2007

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 get 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 two anonymous referees, the editor of this Journal, 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. The second author gratefully acknowledges funding from the Jan Wallander and Tom Hedelius Research Foundation. Addresses for correspondance: javier.ortega@univ-tlse1.fr; thomas.tangeras@ifn.se.

1 Introduction

In 2000, half of the countries in the world had at least one language minority corresponding to more than 10% of their population. This language diversity has recently brought language policies to the forefront of political debate in countries like Malaysia, the ex-Soviet States, Spain, Belgium or the U.S. 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) argues that "for language spread, schools have long been the major formal (organized) mechanisms involved..."(p.116). In other terms, languages which are not given the status of medium of instruction in school tend to be replaced by the 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, around 60% of those living in France did actually not speak French (Grégoire, 1794). Nowadays, everybody 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, which 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 choice, we set-up a model for understanding why some multilingual countries choose unilingual education while others maintain language diversity. We consider an economy with two language groups initially unable to communicate. Value is generated from bilateral production after schooling, among agents speaking the same language. Schooling is a "bundle", as it simultaneously enhances the productivity (or earnings) of students and can modify their language endowment.[2] Schools can be unilingual in either language or bilingual. Individuals attending a unilingual school end up speaking the language of instruction only, while 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 to attend school or not. The interaction among individual school attendance decisions is generated by the network externalities arising from the requirement that production

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

[2] The positive effect of education on earnings is a well established fact in the literature (see e.g. Card, 1999) while the choice of the 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.

partners speak a common language.

While 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 two following systems restricting the use of the language of the dominated group: (i) a unilingual system where only the mother tongue of the dominant is allowed in education, and (ii) an (asymmetric) bilingual system with bilingual schools for the dominated and unilingual schools in their own mother tongue for the dominant.

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 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 have them carry the cost of inter-group 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, since an uneducated member of the dominated group loses the ability to communicate with the members of her own group who take 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 since 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 force them to systematically carry the cost of inter-group 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 due to the technological superiority of bilingual schools. If the central planner can choose the education system, but school attendance remains in the hands of the individuals, laissez-faire is not necessarily optimal anymore, as 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, i.e. we analyse the circumstances, if any, under which the political decision process leads to the adoption of the 'wrong' education system. From a welfare viewpoint, cost efficient communication implies that the minority learns 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. As the dominant group and the majority are the same under majority rule (democracy), this system is shown to always lead to the adoption of a socially optimal decentralised system. Instead, when the dominant are in minority (autocracy), restrictions to the use of the dominated group language are too often implemented.

The basic model is then extended to consider in turn 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, while 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] Like 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, unlike these papers, we consider an investment decision that ties skill acquisition and language acquisition. While in Lazear (1999) agents behave competitively, 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 contributions because the type of network externalities under consideration is endogeneised here, as 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, while John and Yi (2001) provides an explanation based on geography and on inter-generational 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 sizes $L = \{M, N\}$, respectively. We also denote by $-l$ the language group other than l. Political power is in the hands of language group m, whether or not $M \geq 1/2$. For this reason, the ms are also referred to throughout the paper as the

[3] There are 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).

[4] The economics of networks has been extensively studied in the industrial organisation literature, see Farrell and Klemperer (2006) 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.

"dominant" group and the ns as the "dominated" group.[5] Initially, the ms speak $m$ish and the ns speak $n$ish. 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 of producing 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 $\sigma > 0$, when 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:

$$U^b = -c + 1 + \sigma, \tag{1}$$

i.e. 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, $l$ish is the unique language of instruction. We assume that anyone undertaking unilingual $l$ish education ends up speaking $l$ish only. Thus for instance a native $n$ish speaker who attends a unilingual m school learns $m$ish and loses her initial language.[6] Let μ_l^{lu} (respectively μ_l^{-lu}) be the fraction of native $l$ish speakers who attend an l unilingual (respectively $-l$ unilingual) school. Similarly, μ_l^b denotes the proportion of native $l$ish speakers who attend a bilingual school. Finally, $\boldsymbol{\mu}$ is the vector of education levels. The expected utility of attending a unilingual $l \in \{m, n\}$ school is:

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

[5] The 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 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.

[6] Indeed, as shown by linguists (see e.g. Fishman, 1977, for English, and Hagège, 1996, for the case of France) one crucial factor behind language shift in populations over generations is the choice of the language(s) of instruction in school. In other terms, 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 or parental choices are other important factors behind language shift. For a dynamic set-up in which the language spoken by the children (exogenously) depends on the language spoken by the parents and the language spoken in the geographical location, see John and Yi (2001).

The interpretation is that an individual pays c, becomes skilled, and speaks $l$ish when leaving school. She gets $1+\sigma$ from production with the $L(1-\mu_l^{-lu})$ native $l$ish speakers who have not attended a unilingual school in the other language and with the $(1-L)(\mu_{-l}^b+\mu_{-l}^{lu})$ native speakers of the other language that have learnt $l$ish, 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 $l$ish school, as she speaks $l$ish. This individual saves on the cost of education, but gets a value of 1 when producing:

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

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 below 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 thus any individual attending school ends up speaking $m$ish 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. 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, every individual who invests in education will choose a bilingual school, which provides her with a second language and thus *ceteris paribus* enlarges her set of production partners. Then, under laissez faire, denoted by d,[8] $\mu_l^{mu}=\mu_l^{nu}=0$

[7] There 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.

[8] Standing for "deregulated".

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:

$$\triangle U_l^d(\mu_{-l}^b) = -c + (1-L)(1-\mu_{-l}^b) + \sigma \quad \text{for } (l, L) = \{(m, M), (n, N)\}, \tag{4}$$

i.e. a native $l$ish speaker who undertakes education pays c, learns language $-l$ and thus gains as production partners the $(1-L)(1-\mu_{-l}^b)$ native $-l$ speakers who do not speak $l$ish as they do not go to school. This is the communication effect of education. Moreover, there is a productivity gain σ from education, coming from the ability for the bilingual individual to produce an additional amount σ with any other individual.[9]

The Nash equilibria $(\gamma_n^d, \gamma_m^d) = (\gamma_n^b, \gamma_m^b)$ of this game are depicted in Figure 1:

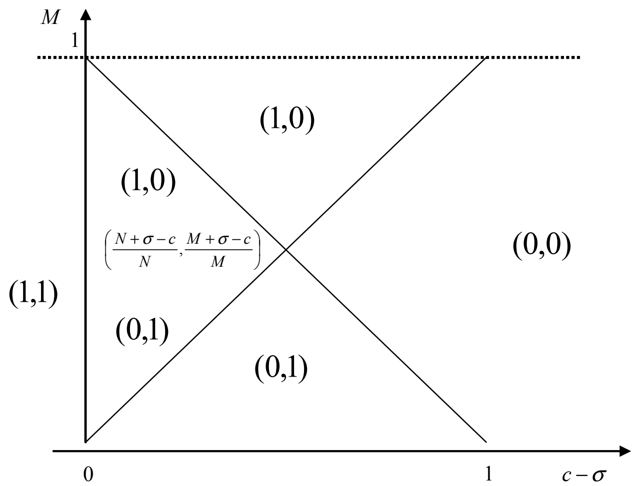

Figure 1 : Equilibrium education levels (γ_n^b, γ_m^b) under laissez-faire

When the productivity gain covers the cost of education ($\sigma > c$), undertaking education is a dominating strategy for everybody. In the other extreme, if education is very expensive ($c > \sigma + \max\{M, N\}$), any educational investment is a dominated strategy, and nobody attends school.

For intermediate values of $c - \sigma$, inter-group 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 educational 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 $-l$ gets fully educated, all its members become bilingual, and can then be reached by the uneducated ls. In this case, education simply has an impact on the productivity of the ls ($\triangle U_l^d(1) = \sigma - c$ from 4), and is unprofitable for them as $c > \sigma$. In turn, the absence of education

[9] In countries where the language of the dominated has low prestige or language groups have (strong) political identities, the dominant may not carefully learn 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, laissez-faire becomes identical to the bilingual system.

among the ls makes group $-l$ willing to invest in schooling only if this generates a sufficient number L of additional production partners, i.e. if $\triangle U^d_{-l}(0) = L + \sigma - c \geq 0$. Hence, $(1, 0)$ and $(0, 1)$ are equilibria respectively if $c - \sigma \in (0, M)$ and if $c - \sigma \in (0, N)$. The two cases are not mutually exclusive, as both equilibria simultaneously arise if the two groups are roughly equal in size and education is not too expensive.[10]

3.2 The n bilingual system

Under the n bilingual system (also referred to for simplicity as the bilingual system), denoted by bi, the ms can only attend m unilingual schools and the ns bilingual schools.[11] Thus, $\mu^b_m = \mu^{nu}_m = 0$, $\mu^{mu}_n = \mu^{nu}_n = 0$, $\mu^{bi}_m = \mu^{mu}_m$, and $\mu^{bi}_n = \mu^b_n$. Using these expressions, subtracting (3) from (1) for $(l, L) = (n, N)$ yields the net benefit from education for a native $n$ish speaker:

$$\triangle U^{bi}_n = -c + M + \sigma. \tag{5}$$

The individual pays c, reaches M additional partners as she learns $m$ish, and gets additional output σ, as she is now skilled and can produce with everybody.[12]

In turn, subtracting (3) from (2) for $(l, L) = (m, M)$, the net benefit from education for the ms is:

$$\triangle U^{bi}_m(\mu^b_n) = -c + (M + N\mu^b_n)\sigma. \tag{6}$$

As both the uneducated and educated ms speak $m$ish only and produce with the $M + N\mu^b_n$ $m$ish speakers, the impact of education for an m is here confined to the productivity effect σ.[13]

The Nash equilibria $(\gamma^{bi}_n, \gamma^{bi}_m) = (\gamma^b_n, \gamma^{um}_m)$ 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 ns since from (5), $\triangle U^{bi}_n = -c + M + \sigma > 0$. Anticipating the high educational levels of the ns, the skill effect is sufficiently strong to render education profitable also for the ms ($\triangle U^{bi}_m(1) > 0$). 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 ms ($\triangle U^{bi}_m(\mu^b_n) < 0$ $\forall \mu^b_n$). Instead, the ns are still willing to pay for education if and only if the M new production

[10] $\triangle U^d_l(\frac{1-L+\sigma-c}{1-L}) = 0$ implies that there also exists an interior laissez-faire equilibrium in this case $(\frac{N+\sigma-c}{N}, \frac{M+\sigma-c}{M})$.

[11] It would have been equivalent here to assume that the ns are free to choose between the three types of schools, as this would have implied that in equilibrium they choose bilingual schools. However, the two assumptions are not anymore equivalent in section 6.3 where we consider the case of a higher cost of bilingual schools.

[12] Comparing (5) with (4) for $l = n$, it appears that schooling incentives are higher here for the ns than under laissez-faire, since the ms never learn $n$ish, and thus schooling always enlarges the set of production partners for the ns. Instead, under laissez-faire, an uneducated n could still produce with the $M\mu^d_m$ bilingual ms.

[13] This implies that the representative m has lower incentives for education than under laissez-faire (see 4), where education could enlarge her set of production partners.

possibilities it generates compensates for the insufficient productivity effect $c-\sigma$. Hence, the ns become bilingual provided $c-\sigma \in (0, M)$ and otherwise remain uneducated.[14]

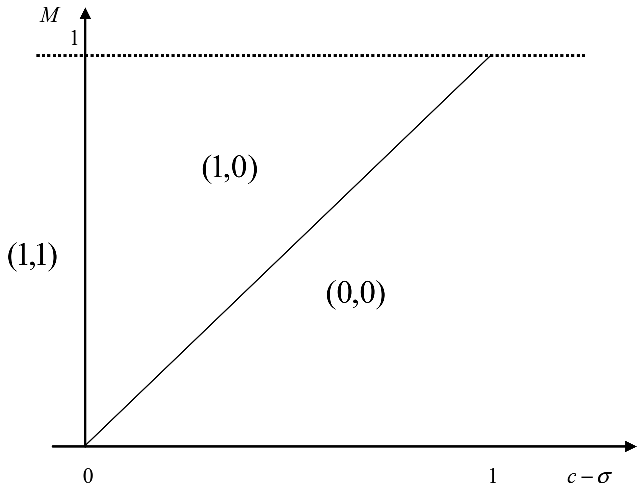

Figure 2 : Equilibrium education levels $(\gamma_n^b, \gamma_m^{um})$ under bilingualism

3.3 The m unilingual system

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

$$\triangle U_n^{uni}(\mu_n^{mu}) = -c + (M + N\mu_n^{mu})\sigma + M + N\mu_n^{mu} - N(1-\mu_n^{mu}). \tag{7}$$

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

Equation (7) generates an insight that is crucial to the understanding of the preferences over education systems. For the ns, 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 $n$ish speakers and in particular the larger the number of other ns attending the unilingual school. This positive communication externality is thus at the origin of a *bandwagon*

[14] Comparing Figure 2 with Figure 1, the ns take more education than under laissez-faire, while the reverse is true for the ms.

or *snowball* effect in the schooling decisions of the ns. Indeed, it is easy to check from (7) that the net return to schooling for an n is increasing in the number of ns 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 ns take education, an interior equilibrium may exist.[15] Intuitively, for the bandwagon effect to play a role in equilibrium, the dominated group must be sufficiently large, for otherwise avoiding to go to school and restricting to intra-group production is never profitable for the ns.

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 (i) for full details). The $\overline{M}(c-\sigma)$ line characterises the critical size of the dominant group below which the bandwagon effect comes into play. It is upward sloping since 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 $\overline{M}(c-\sigma)$ line, the dominant group is so large and thus the productivity gain and communication effect so strong relative to the cost of education, that education is a dominant strategy for the ns. Then, the ms get educated if schooling is sufficiently cheap ($c < \sigma$), and abstain from education otherwise.

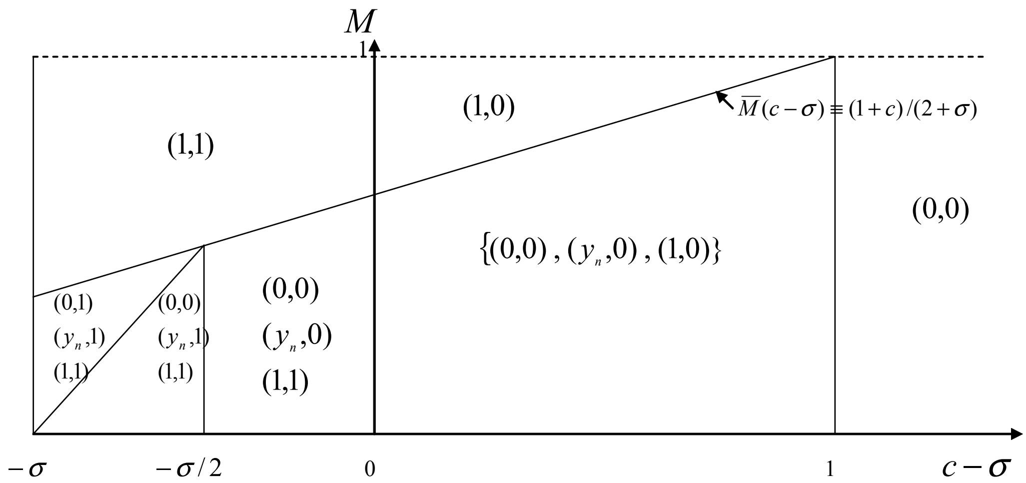

Figure 3: Equilibrium education levels $(\gamma_n^{mu}, \gamma_m^{mu})$ under the unilingual system, with $y_n \equiv (\overline{M} - M)/N$

South-east of the $\overline{M}$ line 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 ns *de facto* discourages education.

4 Welfare

Expected welfare is obtained by adding up individual utility levels:

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

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

In the presence of a benevolent social planner able to enforce welfare maximising education levels under each system, the following proposition can be stated:

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

Proof. see appendix (ii). ■

Given identical costs of unilingual and bilingual schools, the total expenditures associated to a given educational level are the same independently of the education system. Instead, production (and thus expected utility) is larger the larger the share of bilingual individuals. As 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 students spend in their studies, even in a system with mandatory education.[16] To capture this degree of freedom, consider therefore a situation in which the central planner picks the educational system, and individuals decide whether to attend school. Under decentralised school attendance choice, the following proposition can be stated:

Proposition 2 *In the choice between laissez-faire, bilingualism, and unilingualism, the system yielding a higher decentralised welfare is depicted in Figure 4.*[17]

Proof. see appendix (ii). ■

Figure 4 shows that laissez-faire ceases to be optimal in regions (I) and (IIa) due 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 gets educated. Given $M > 0.5$, only the ns should attend school. However, under laissez-faire, the schooling incentives of the ms are strong if few of the ns are expected to take education. The majority (i.e the wrong group) may therefore end up undertaking education. In contrast, under the two other systems, the ms cannot learn $n$ish, which reduces their incentives to undertake education, and guarantees that only the ns, 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 since the bandwagon effect that 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, now because it fails to generate sufficient schooling incentives.

[16] There 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 references therein).

[17] The figure is drawn assuming $\sigma < 1$. This 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 that are under consideration. The ranking in this proposition builds on the exclusion of interior unstable equilibria in parts of the parameter range for which $M < 0.5$. All subsequent results hold for comparisons of all equilibria.

Instead, the bandwagon effect can raise schooling under the unilingual system to a level that would be impossible to reach under any of the two other systems. Thus, despite the fact that bilingual schools are technologically superior, a system which restricts (some) individuals from access to bilingual schools may be better because it provides stronger schooling incentives than laissez-faire.

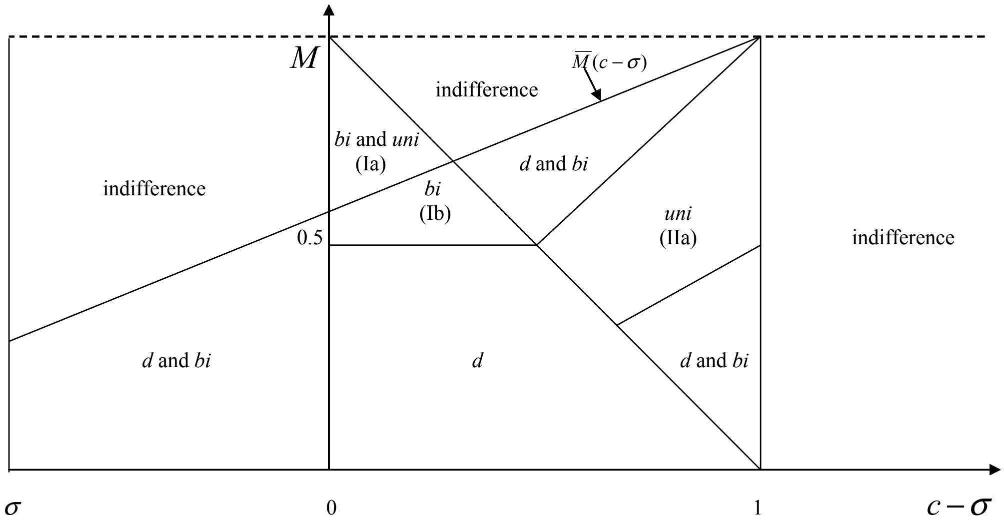

Figure 4: System yielding a higher decentralised welfare: comparison among laissez-faire, unilingualism, and bilingualism.

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

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

Proof. see appendix (ii). ■

When $M > 0.5$, 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, but as 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 thus all systems are equivalent.

The three systems may all be suboptimal whenever the dominant group is in minority. For example, the m bilingual system (the ms become bilingual, the ns 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 $M \in (c - \sigma, 0.5)$ and $N > c - \sigma$ when neither of the three systems above is capable of doing so.

5 The choice of education system

This section analyses how the ms and the ns 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 set-up, and whether this conflict is of the expected type, i.e. a situation where the dominated favour the use of $n$ish in education while the dominant oppose to it. In addition, we determine which political rules, if any, enable society to reach a decentralised optimum.

The ms 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 ms prefer to laissez-faire, the unilingual and the bilingual systems, and determines when each of these three systems is chosen:

Proposition 4 *In decentralised equilibrium, (i) if the cost of education is low ($c < \sigma$) or high ($c > 1 + \sigma$), 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 ($M > c - \sigma$), the dominant weakly prefer bilingualism to any other system, or are indifferent; (b) if the dominant group is sufficiently small ($M < c - \sigma$), there exists no education system that the dominant prefer to the unilingual system.*

Proof. see appendix (ii). ■

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 ns. When $c < \sigma$, education is profitable also for the ms, since the cost is always redeemed by increased productivity. Instead, for instance, unilingualism fails to guarantee that all the ns attend school and thus the ms, who would not learn $n$ish, may not be able to produce with everybody, and do not choose that system.

When education gets more expensive ($\sigma < c < 1 + \sigma$), the maximum gain from education σ becomes insufficient to cover the cost of education, and thus the ms try to avoid undertaking education, while still producing with everybody. Laissez-faire is not always useful for this purpose when the dominant group is large ($M > c - \sigma$). Indeed, under this system, if the ns choose not to attend school, 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 themselves from learning $n$ish. With this commitment device, inter-group production can only be initiated by the ns, which in turn lowers (raises) the education incentives of the ms (ns) and guarantees maximum communication paid for by the dominated. When the dominant group is small, i.e. $M < c - \sigma$, the bandwagon effect may generate the best outcome for the ms under unilingualism, while laissez-faire and bilingualism fail to create sufficient incentives for the ns to undertake education.[18] More

[18] For $c - \sigma < 1 - M$ in this area, the choice of the ms between laissez-faire and unilingualism is indeterminate.

generally, it can be shown that any other system has an equilibrium where the ns do not invest in education, and thus cannot be better than unilingualism from the ms' viewpoint.[19]

Proposition 5 in turn characterises the preferences of the dominated over 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 (ii). ■

Under laissez-faire, the ms learn $n$ish whenever they undertake education, and thus the dominated can gain production partners without investing in education, which explains why they prefer laissez-faire over bilingualism. Instead, bilingualism and unilingualism have in common that the cost of inter-group communication is always borne by the ns. And between these two, the ns prefer bilingualism due to the absence of negative network externalities and thus of bandwagon effects. In contrast, under unilingualism, the ns are in some cases "compelled" to undertake education even if this is relatively expensive, since this is the only way of learning $m$ish and avoiding to remain insulated from the rest of society.

The joint implication of propositions 4 and 5 is that language conflict always comes with the desire of the ms to restrict the use of $n$ish in education, and the opposition of the ns to this choice (see also Figure 5). This is an interesting result since language conflict is here of the expected type and does not rely on any direct utility enjoyed by the agents from speaking their own native language.[20] In contrast, language conflict is an equilibrium phenomenon. The dominated want their language to be a means of instruction in school not because they "like it" but rather because abandoning it would force them to overinvest in education. In the same way, by legally restricting the amount of n speakers, the dominant maximise the incentive for learning $m$ish, thereby forcing the cost of communication on the dominated. In other terms, the ms free-ride on the costs of speaking a common language, which are entirely borne by the ns.

More precisely, in region (I) of Figure 5, the dominant support a regulation that prevents themselves from learning $n$ish. Intuitively, the ms need a way of committing not to learn $n$ish for the ns to pay for education. This is not possible under laissez-faire, since it might be individually rational for the ms to become bilingual.[21] In contrast, the ns would like the ms to learn $n$ish so as to get additional production partners without paying for education, as is achieved under

While the best outcome can only be reached under unilingualism, the ns prefer $(\gamma_n^d, \gamma_m^d) = (0,1)$ to $(\gamma_n^{uni}, \gamma_m^{uni}) = (0,0)$.

[19] Finally, when education is very costly ($c > 1 + \sigma$) the choice of system does not matter since nobody ever goes to school, regardless of the system.

[20] Adding an exogenous utility term of speaking one's mothertongue 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, as they maintain their mother tongue under all three systems.

[21] In region (Ic), the ms prefer bilingualism to unilingualism to eliminate possible negative bandwagon effect under unilingualism. Instead, the bandwagon effect does not play in region (Ia), which explains why the ms are indifferent between the two systems.

laissez-faire. In region (II), the dominant want to ban the use of $n$ish in education *for everybody.* As explained above, unilingualism in this region provides stronger incentives than bilingualism or laissez-faire for the dominated to become educated. Then, the ms go for the unilingual system, hoping to lock in the ns in a high-education equilibrium, whereas the dominated go for bilingualism or laissez-faire, precisely in order to avoid the same situation.

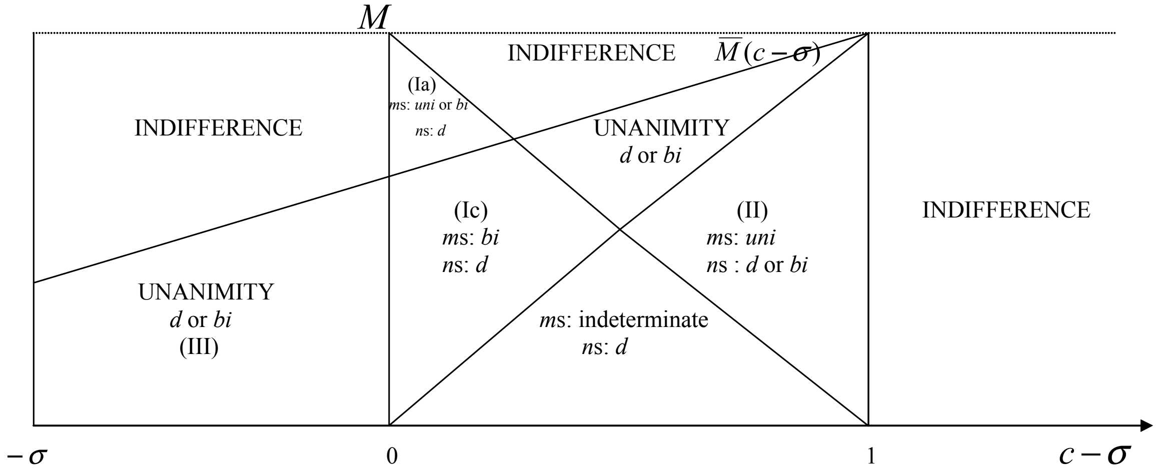

Figure 5: Choice of education system

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 takes education, and the choice of system does not matter.

Having identified the preferences of the two groups, we are now able to discuss the welfare properties of various political systems regarding the choice of language of instruction. Recall that the education system is chosen to maximise the expected utility of the dominant (the ms) independently on whether they are a majority. Majority rule, which can be interpreted as a democratic system, thus corresponds to a situation where $M > 0.5$, whereas minority rule (i.e. autocracy) prevails whenever $M < 0.5$ Comparing Figures 4 and 5, and using Proposition 3:

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

The dominant achieve maximum communication by maximising the number of $m$ish speakers. From a welfare viewpoint, cost efficient communication implies that the minority learns the majority language. The dominant and the majority are the same under majority rule, which explains why majority rule works well. Under minority rule, cost efficient communication requires that the minority (the ms) learn $n$ish but the ms impose systems which restrict access to n unilingual schools or to bilingual schools and result in the majority getting educated.

6 Extensions

This section extends the basic model to consider in turn (i) education subsidies, (ii) the possibility for native $l$ish speakers attending a $-l$ unilingual school to retain their mother tongue and thus become bilingual, (iii) an open economy with F_l $l$ish speakers abroad, $l = \{m, n\}$, (iv) an additional cost κ of attending bilingual schools, (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 (iii), while the full analytical details of all the extensions are in a Technical Appendix.[22]

6.1 Education subsidies

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

Proposition 7 *There exists an education subsidy targeted to the minority, 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), 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 which selects among strict equilibria.

[22] The Technical Appendix is available upon request and at http://www.ifn.se/thomast.

6.2 Mother tongue persistence

Imagine that mother tongue is retained with probability α after unilingual education in the other language. Then, every n attending a unilingual $m$ish school under unilingualism becomes bilingual with probability α and is thus subject to the bandwagon effect with probability $1 - \alpha$ only.[23] This results in a weaker bandwagon effect under unilingualism and in turn 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 ms prefer unilingualism, as well as an area in which unilingualism is a decentralised optimal system. In addition, 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 $l$ish speakers abroad, $l = \{m, n\}$. As the foreigners never shift language nor 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 speaker 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 can be stated:

Proposition 8 *If $F_n \geq \max\{M + F_m; (1 + F_m)/(2 + \sigma)\}$, 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 ms choose the only system which allows them to learn $n$ish and thus reach the large number of $n$ish speaking foreigners, while the ns still have no incentive to restrict the use of their own language.

A second new result arises when F_n is large relative to N ($F_n > \frac{N}{1+\sigma}$). In this case, any native $n$ish speaker keeps 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 cannot anymore generate high schooling levels among the ns and ceases to be interesting to the eyes of 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 N, most of the qualitative results of the baseline model still hold, as in this case opening the

[23] Nobody ever attends at equilibrium a unilingual school in the other language under laissez-faire or bilingualism, which implies that these two systems remain unchanged.

economy just enlarges its size. In particular, unilingualism and bilingualism are chosen by the ms 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.[24]

6.4 Higher cost of bilingual schools

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

While the central planner never chooses the bilingual system due to $\kappa > 0$, the ms may still implement bilingualism since the extra cost of education falls on the ns. Excessive bilingualism occurs whenever the two groups are roughly equal in size and $M < 0.5$. 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 instead of in 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 question the robustness of the baseline results. However, bilingual school attendance under laissez-faire is restored in an open economy, provided that the differential cost κ of becoming bilingual is compensated by a larger expansion in the set of production partners abroad. In this case, bilingual education may be again 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, despite the fact that French speakers were a minority.[25] In Flanders, Dutch was partly used as language of instruction in primary schools, but 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

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

[25] Dutch (Flemish) speakers accounted for 57% of the population in 1846 (McRae, 1986).

overwhelmingly French-speaking, even in the Flemish provinces" (McRae, 1986, p.21) and only 46,000 electors 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 issue in the political choices during the French Revolution (1789-1794), with a series of French-unilingual decrees 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 the French language policy (Weber, 1976). In 1794, only roughly 40 percent of the population were native French-speakers (Calvet, 2002, p. 218).[26] Among the other language groups, the biggest was Occitan, and next came Breton and Alsacian.[27]

Our model predicts that minority size may not be the most relevant variable for understanding the choice of education system. Indeed, from inspection of Figure 5, in the absence of information concerning the net cost of education $c - \sigma$, the size 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 French only in each of the 89 *départements*, together with the proportion of French-speakers in the local population. The data show that there was regional variation in educational systems at that time, before the introduction of the Ferry Laws in 1880-82, 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 department-level variables for the 89 departments. 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 department. This may indicate that parents were willing to invest more in education if schools were in French, most likely following a social mobility argument, as French was necessary in skilled occupations.

This first regression however does not take into account that in the 55 fully French-speaking departments the possibility of having non-French speaking schools was not even considered. We control for this introducing a dummy variable for unilingual French-speaking departments in the regressions of Columns 2 and 3. The results 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.

[26] This estimate is based on the language report conducted by Grégoire (1794).

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

Table 1: The proportion of French-unilingual public schools at the department level in France (1863)

Dependent variable: Proportion French-unilingual public schools			
French-speakers in the population	1.17**	.805	.210
	(.442)	(.679)	(.753)
Amount paid for education by a family	.318***	.291**	.267**
	(.108)	(.117)	(.117)
Log income per head	.960*		1.2
	(.482)		(1.24)
Unilingual French-speaking department (dummy)		3.03***	2.97***
		(.584)	(.795)

Notes: The figures reported are the coefficients obtained from tobit estimation. Standard errors in parentheses. *, ** and *** denote significance at 10%, at 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). The remaining data are from Ministère de l'Instruction Publique (1878). All data refer to 1863.

7.3 Finland

The current institutional language framework in Finland was set up with the Constitution of 1919 and a series of language laws, the most important being approved in 1922. In 1920, the Swedish-speakers constituted only 11% of the population (McRae, 1997), the rest of the population being Finnish-speaking, except for a tiny Sami group. Nevertheless, the Constitution recognises Finnish and Swedish as national languages on an equal basis. The language clauses of the Constitution were approved by very large majorities, ranging from 88% to 96% (see 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, but rather the proposed symmetric bilingual system. For this reason, we argue that the Finnish-speakers were at that time in terms of our model the dominant group, despite the fact that Swedish speakers had been historically over-represented among the elites.

The educational system is such that each municipality needs to provide schooling in the minority language (Swedish or Finnish) when a minimum number of parents requires it. Given the symmetry of the system, the unanimity for bilingual schools in Finland may be delivered in our model as unanimity for laissez-faire, and can be explained in an economy with cross-border spillovers in the light of Proposition 8. According to this proposition, if the language distribution abroad is skewed towards the language of the dominated (i.e. here the Swedish-speakers), laissez-faire is always unanimously chosen. Taking into account the number of Finnish and Swedish speakers in Finland and Sweden in 1920, $M = 0.89$, $N = 0.11$, $F_m = 0.0097$ and $F_n = 1.89$, so the condition on the skewness of the language distribution in Proposition 8 is satisfied even for

$\sigma = 0$.[28]

8 Conclusion and further discussion

While 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 coming to decide the language of instruction in school, nor has it always been the case that both groups have agreed upon a unilingual or a bilingual system. A possible way to go for understanding this variety of situations is to assume that agents enjoy some utility from speaking their own language and that compromise over language issues through political bargaining is only reached in some cases.

Here, we take a different stand on the issue and propose a model centered on the 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. Consequently, the dominated either bear a large part of the costs necessary to the adoption of 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 endogenously arise as the result of an economic conflict.

Should we expect language conflicts to end up 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, while the use of the dominated group language is too often restricted in autocracies.

While our model delivers a number of empirically grounded results using a parsimonious set of economic and political features, there are aspects of existing language policies that can only be rationalised with further modeling assumptions. In particular, our model does not address the within country regional distribution of languages, nor does it allow for a two-tier political decision process, say at the local and federal level. Some countries admit multiple languages of instruction, while restricting the 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 a possible impact of the newly declared freedom of movements inside the country on the language composition of

[28] 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 (Kungliga Statistiska Centralbyrån, 1924, p.14).

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 still in work today (except for a greater role given to Romansh), part of the cantons with sizeable language minorities implement unilingual education systems. In Belgium, a similar system has been instituted from the 1962-63 language laws, and currently Dutch is the unique language of instruction in Flanders, and the same applies for French in Wallonia while French and Dutch co-exist as languages of instruction in the Brussels-capital region.[29] A framework that accounts for such institutional choices should include geographical distance and the cross-regional migration decision 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). This was however overturned by a new active bilingual policy at the federal level (Canada Constitution Act, 1982), establishing 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

(i) Equilibrium education levels under unilingualism

(i) Education levels of the ns. Define $\overline{M} \equiv \frac{1+c}{2+\sigma}$, substitute into (7) and rewrite to get $\triangle U_n^{uni}(\mu_n^{um}) = (2+\sigma)\left(N\mu_n^{um} + M - \overline{M}\right)$. $c > 1+\sigma$, implies $\overline{M} > 1$, hence $\triangle U_n^{uni}(\mu_n^{um}) < 0$ and $\gamma_n^{uni} = 0$ in this case. Next, $c < 1+\sigma$ implies $\overline{M} < 1$ and education is a strictly dominating strategy $\forall\ M > \overline{M}$. Hence, $\gamma_n^{uni} = 1$ in this case. For $c < 1+\sigma$ and $M < \overline{M}$, $\triangle U_n^{uni}(1) = (2+\sigma)\left(1-\overline{M}\right) > 0$, $\triangle U_n^{uni}(0) = (2+\sigma)\left(M-\overline{M}\right) < 0$ and $\triangle U_n^{uni}((\overline{M}-M)/N) = 0$ imply that all $\gamma_n^{uni} \in \{0, (\overline{M}-M)/N, 1\}$ are equilibria. (ii) Education levels of the ms. From (6), $\triangle U_m^{uni}(\mu_n^{um}) \leq \triangle U_m^{uni}(1) = \sigma - c$ implies that $\gamma_m^{uni} = 0$ is the unique equilibrium for $c > \sigma$. Next, $\triangle U_m^{uni}(\mu_n^{um}) \geq \triangle U_m^{uni}(0) = M\sigma - c$ implies $\gamma_m^{uni} = 1\ \forall\ M > c/\sigma$. For $c < \sigma$ and $M < c/\sigma$, the equilibrium education level γ_m^{uni} depends on γ_n^{uni}. First, $\gamma_n^{uni} = 1$ implies $\gamma_m^{uni} = 1$ since $\triangle U_m^{uni}(1) = \sigma - c > 0$. Second, $\gamma_n^{uni} = 0$ implies $\gamma_m^{uni} = 0$ since $\triangle U_m^{uni}(0) = (M - c/\sigma)\sigma < 0$. Finally, $\triangle U_m^{uni}\left(\frac{\overline{M}-M}{N}\right) = \overline{M}\sigma - c = \frac{\sigma-2c}{2+\sigma}$. Hence, $\gamma_n^{uni} = (\overline{M}-M)/N$ implies $\gamma_m^{uni} = 1$ for $c < \sigma/2$ and $\gamma_n^{uni} = 0$ for $c > \sigma/2$.

[29] There are some 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.

(ii) Proofs

Let Γ^s be the set of equilibria under education system s and $\boldsymbol{\gamma}^s \in \Gamma^s$ a specific equilibrium under s. Let $u_l^s(\boldsymbol{\gamma}^s) = \gamma_l^{mu} U^{mu}(\boldsymbol{\gamma}^s) + \gamma_l^{nu} U^{nu}(\boldsymbol{\gamma}^s) + \gamma_l^b U^b + (1 - \gamma_l^b - \gamma_l^{mu} - \gamma_l^{nu})\underline{U}_l(\boldsymbol{\gamma}^s)$ be the indirect utility for a native $l$ish speaker under s given the play of $\boldsymbol{\gamma}^s$. 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 $\min\{u_l^s(\boldsymbol{\gamma}^s)|\boldsymbol{\gamma}^s \in \Gamma^s\} \geq \max\{u_l^{s'}(\boldsymbol{\gamma}^{s'})|\boldsymbol{\gamma}^{s'} \in \Gamma^{s'}\}$. 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 $l$ish speaker is indifferent between the two $(s \sim_l s')$ if all equilibria generate $u_l^s(\boldsymbol{\gamma}^s) = u_l^{s'}(\boldsymbol{\gamma}^{s'})$. Subscript w on the preference ordering denotes the social welfare ordering. $W(\boldsymbol{\gamma}^s)$ denotes expected welfare, as defined in (8) for equilibrium $\boldsymbol{\gamma}^s$. If $s \in \{d, uni, bi\}$, 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.

Proposition 1 For any distribution $\boldsymbol{\mu}$ of education levels, we show that there exists an alternative distribution $\widetilde{\boldsymbol{\mu}}$ under d, such that $W(\widetilde{\boldsymbol{\mu}}) \geq W(\boldsymbol{\mu})$. Pick any $\boldsymbol{\mu}$, and consider $\widetilde{\boldsymbol{\mu}}$ where a fraction $\widetilde{\mu}_l = \mu_l = \mu_l^b + \mu_l^{mu} + \mu_l^{nu}$ of the ls $(l \in \{m, n\})$ are forced to take education, but free to choose the type of school. Naturally, everybody chooses a bilingual school under d. Hence, $\widetilde{\mu}_l^{mu} = \widetilde{\mu}_l^{nu} = 0$ and $\widetilde{\mu}_l^b = \widetilde{\mu}_l = \mu_l \; \forall \; l \in \{m, n\}$. The welfare difference is

$$W(\widetilde{\boldsymbol{\mu}}) - W(\boldsymbol{\mu}) = \textstyle\sum_{(l,L)=\{(m,M),(n,N)\}}(M\mu_m^{lu} + N\mu_n^{lu})(U^b - U^{lu}(\boldsymbol{\mu})) + L(1-\mu_l)(\underline{U}_l(\widetilde{\boldsymbol{\mu}}) - \underline{U}_l(\boldsymbol{\mu})).$$

$U^b \geq U^{lu}(\boldsymbol{\mu})$ and $\underline{U}_l(\widetilde{\boldsymbol{\mu}}) - \underline{U}_l(\boldsymbol{\mu}) = M\mu_m^{-lu} + N\mu_n^{-lu} \geq 0$ for $l \in \{m, n\}$ imply $W(\widetilde{\boldsymbol{\mu}}) \geq W(\boldsymbol{\mu})$. Determination of $\mathbf{x} = (x_n^d, x_m^d)$: plugging $\mu_l^{mu} = \mu_l^{nu} = 0$ and $\mu_l^b = \mu_l^d$ for $l \in \{m, n\}$ into (8) yields laissez-faire welfare $W^d(\mu_n^d, \mu_m^d) = 1 - 2MN(1-\mu_m^d)(1-\mu_n^d) + (\sigma - c)\left(M\mu_m^d + N\mu_n^d\right)$. Then, $W^d(1,1) - W^d(\mu_n^d, \mu_m^d) = 2MN(1-\mu_m^d)(1-\mu_n^d) + (\sigma - c)\left(M(1-\mu_m^d) + N(1-\mu_n^d)\right)$ implies $\mathbf{x} = (1,1) \; \forall \; c \leq \sigma$. Similarly, $W^d(1,0) - W^d(\mu_n^d, \mu_m^d) = \mu_m^d(c-\sigma)(2M - 1 + N\mu_n^d) + N(2M + \sigma - c)(1-\mu_m^d)(1-\mu_n^d)$ implies $\mathbf{x} = (1,0) \; \forall \; c - \sigma \in (0{,}2M)$ and $M > 0.5$. The proof that $\mathbf{x} = (0,1)$ if $c - \sigma \in (0{,}2N)$ and $N > 0.5$ is analogous. Finally, $W^d(0,0) - W^d(\mu_n^d, \mu_m^d) = M\mu_m^d(c - \sigma - 2N) + N\mu_n^d(c - \sigma - 2M) + 2MN\mu_m^d\mu_n^d$ implies $\mathbf{x} = (0,0) \; \forall \; c - \sigma > 2\max\{M; N\}$. ■

Proposition 2 The equilibrium education levels $\gamma^s \epsilon \; \Gamma^s$ for $s \in \{d, uni, bi\}$ are taken from Section 3 and $\mathbf{x}$ from the preceding proof. (i) $c < \sigma$: d and bi both uniquely implement maximal communication and education, and so does uni provided also $M > \overline{M} \equiv \frac{1+c}{2+\sigma}$. Hence, $W^d = W^{bi} = W^{uni} = W^d(\mathbf{x})$ in this case. For $M < \overline{M}$, uni in addition generates undereducation equilibria, hence ceases to be optimal. (ii) $c > 1 + \sigma$: no system generates positive education, hence all systems are equivalent. (iiia) $\sigma < c < 1 + \sigma$ and $M > 0.5$: If $M > \overline{M}$ and $N < c - \sigma$ in addition hold, all three systems uniquely implement maximal communication through maximal (no) education of the minority (majority) in the majority language, i.e. all systems attain $W^d(\mathbf{x})$.

If instead, $M > \overline{M}$, but $N > c-\sigma$ (region Ia) d may educate the wrong group, hence is no longer optimal. For, $N < c-\sigma$ and $M \in (c-\sigma, \overline{M})$, *uni* may lead to undereducation, hence is no longer optimal. Fourth, if $M \in (1/2, \overline{M})$ and $N > c-\sigma$, (region Ib) only *bi* of all the three systems uniquely attain $W^d(\mathbf{x})$. Finally, if $M \in (1/2, c-\sigma)$, (a subset of region II) $\Gamma^{bi} = \Gamma^d = (0,0)$, while $\Gamma^{uni} = \{(1,0); (\frac{\overline{M}-M}{N}, 0); (0,0)\}$. Since $W^{uni}(1,0) = W^d(\mathbf{x})$, $W^{uni}(\frac{\overline{M}-M}{N}, 0) - W^{uni}(0,0) = (M-N)(\overline{M}-M) > 0$ and $W^{uni}(0,0) = W^{bi}(0,0) = W^d(0,0)$, $uni \succeq_w bi \sim_w d$. (iiib) $\sigma < c < 1+\sigma$ and $M < 0.5$: d is the only system of the three which can possibly implement $\mathbf{x}$ as an equilibrium. This happens for $N > c-\sigma$. If, in addition, $M < c-\sigma$, d uniquely implements $\mathbf{x}$, but if $M > c-\sigma$, 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 $W^s(1,0) - W^s(0,0) = N(2M+\sigma-c) > 0$ for $s = uni, bi$ implies $d \succeq_w uni$ and $d \succeq_w bi$. Next, $\Gamma^{bi} = \Gamma^d = \{0,0\}$ if $N < c-\sigma$. $W^{bi}(0,0) > W^{uni}(1,0) \Leftrightarrow M < \frac{c-\sigma}{2}$ implies that $d \sim_w bi \succeq_w uni$ for $M < (c-\sigma)/2$ while $uni \succeq_w bi \sim_w d$ for $M > (c-\sigma)/2$ considering stable equilibria only (we are then back to region II). ■

Proposition 3 The preceding proof shows that the systems depicted in Figure 4 all attain $W^d(\mathbf{x})$ through a unique equilibrium whenever $M > 0.5$ and $c < 1+\sigma$, except for $M \in (1/2, c-\sigma)$. Thus, they cannot be outperformed by any other system. For $M \in (1/2, c-\sigma)$, $W^{uni}(1,0) = W^d(\mathbf{x})$, implying that any system that outperforms *uni* must uniquely implement $W^d(\mathbf{x})$. However, as long as $M < c-\sigma$, any system will necessarily have $\widetilde{\gamma}_n = \widetilde{\gamma}_n^b + \widetilde{\gamma}_n^{mu} + \widetilde{\gamma}_n^{nu} = 0$ as part of a stable equilibrium. Indeed, $\triangle U_n^{lu}(\widetilde{\gamma}) \leq \triangle U_n^b(\widetilde{\gamma})$ always holds (for $l = m, n$) and $\triangle U_n^b(\widetilde{\gamma}) = \sigma + M - c - M(\widetilde{\gamma}_m^b + \widetilde{\gamma}_m^{nu}) < 0$ for $M < c-\sigma$ and $\widetilde{\gamma}_n = 0$. Hence, *uni* is a decentralised optimum in this case. Finally, no system generates positive education for $c > 1+\sigma$, and hence all systems are equivalent in welfare terms in this case. ■

Proposition 4 Compare an equilibrium $\boldsymbol{\gamma}^s$ under system $s \in \{d, uni, bi\}$ such that $\gamma_n^{mu} + \gamma_n^b = 1$ to an arbitrary equilibrium $\widetilde{\boldsymbol{\gamma}}$ under system s', where γ_l^s and $\widetilde{\gamma}_l$ denote the aggregate education of group l in each of the equilibria. Note first that $U_m^{mu}(\boldsymbol{\gamma}^s) = U^b = 1+\sigma-c$ and $\underline{U}_m(\boldsymbol{\gamma}^s) = 1$. After some manipulations, $u_m^s(\boldsymbol{\gamma}^s) - u_m^{s'}(\widetilde{\boldsymbol{\gamma}}) = (\gamma_m^s - \widetilde{\gamma}_m)\triangle U_m^s(1) + \widetilde{\gamma}_m^{nu}(M(1-\widetilde{\gamma}_m^b - \widetilde{\gamma}_m^{nu}) + N\widetilde{\gamma}_n^{mu})$ $(1+\sigma) + (\widetilde{\gamma}_m^{nu}(1+\sigma) + 1 - \widetilde{\gamma}_m)(M\widetilde{\gamma}_m^{nu} + N(1-\widetilde{\gamma}_n^b - \widetilde{\gamma}_n^{mu}))$ which is non-negative since $\gamma_m^s < \widetilde{\gamma}_m \leq 1$ implies $\triangle U_m^s(1) \leq 0$, and $\gamma_m^s > \gamma_m \geq 0$ implies $\triangle U_m^s(1) \geq 0$. Thus $\boldsymbol{\gamma}^s$ 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 d. Hence, the two systems are equivalent. For $c > 1+\sigma$ all systems are characterised by zero education, and are thus equivalent. Part (ii), $\sigma < c < 1+\sigma$: For $M > c-\sigma$, $\gamma_n^b = 1$ is reached as a unique equilibrium under *bi*. For $M < c-\sigma$, $\gamma_n^{mu} = 1$ and zero education in both groups are the two stable unilingual equilibria. In order for a competing system to outperform *uni*, it must have $\widetilde{\gamma}_n^{mu} + \widetilde{\gamma}_n^b = 1$ in every equilibrium for $M < c-\sigma$. However, any education system necessarily has $\widetilde{\gamma}_n = 0$ as a (stable) equilibrium in this interval (see the preceding proof). ■

Proposition 5 Consider arbitrary equilibria $\boldsymbol{\gamma}^d$, $\boldsymbol{\gamma}^{uni}$ and $\boldsymbol{\gamma}^{bi}$. By manipulating terms, $u_n^d(\boldsymbol{\gamma}^d) - u_n^{bi}(\boldsymbol{\gamma}^{bi}) = (\gamma_n^d - \gamma_n^{bi}) \triangle U_n^d(\gamma_m^d) + M\gamma_m^d(1-\gamma_n^{bi}) \geq 0$ and $u_n^{bi}(\boldsymbol{\gamma}^{bi}) - u_n^{uni}(\boldsymbol{\gamma}^{uni}) = (\gamma_n^{bi} - \gamma_n^{uni}) \triangle U_n^{bi}(\gamma^{bi}) + N\gamma_n^{uni}(1-\gamma_n^{uni})(2+\sigma) \geq 0$ which imply $u_n^d(\boldsymbol{\gamma}^d) \geq u_n^{bi}(\boldsymbol{\gamma}^{bi}) \geq u_n^{uni}(\boldsymbol{\gamma}^{uni})$ $\forall \boldsymbol{\gamma}^d, \boldsymbol{\gamma}^{uni}$ and $\boldsymbol{\gamma}^{bi}$. ■

(iii) Extensions

The expected utility for a native $l$ish speaker of attending a bilingual school, a unilingual $l$ish school, and a unilingual school in the other language $(-l)$ are respectively:

$$\begin{aligned}
U^b &= -c - \kappa + (1+F)(1+\sigma), \\
U_l^{lu}(\boldsymbol{\mu}) &= -c + [L(1-(1-\alpha)\mu_l^{-lu}) + (1-L)(\mu_{-l}^b + \mu_{-l}^{lu}) + F_l](1+\sigma), \\
U_l^{-lu}(\boldsymbol{\mu}) &= -c + \alpha(1+F)(1+\sigma) + (1-\alpha)[L(\mu_l^b + \mu_l^{-lu}) + (1-L)(1-(1-\alpha)\mu_{-l}^{lu}) + F_{-l}](1+\sigma),
\end{aligned}$$

while the utility of not taking education for this individual is:

$$\underline{U}_l(\boldsymbol{\mu}) = L(1-(1-\alpha)\mu_l^{-lu}) + (1-L)(\mu_{-l}^b + \mu_{-l}^{lu}) + F_l.$$

References

[1] Barbaud, Philippe (1998) "French in Quebec", in John Edwards (ed.) *Language in Canada.* Cambridge University Press.

[2] 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, No. 2, pp. 229-254.

[3] Encyclopaedia Britannica (2003) *Britannica Book of the Year 2002.*

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

[5] Card, David (1999) "The Causal Effect of Education on Earnings", in Orley Ashenfelter and David Card (eds.) *Handbook of Labor Economics,* vol 3A, chapter 30, pp. 1801-63.

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

[7] Chiswick, Barry R. and Paul W. Miller (1995) "The Endogeneity between Language and Earnings: International Analyses", *Journal of Labour Economics* 13 (April), pp. 246-88.

[8] Church, Jeffrey and Ian King (1993) "Bilingualism and network externalities" *Canadian Journal of Economics-Revue canadienne d'Economique* XXV, No. 2, pp. 337-345.

[9] Eduskunta-Riksdag (1920) Valtiopäivät 1919. Pöytäkirhat I. Istunnot 1 53.

[10] Farrell, Joe and Paul Klemperer (2006) "Coordination and Lock-In: Competition with Switching Costs and Network Effects", in Mark Armstrong and Robert Porter (eds.) *Handbook of Industrial Organization*, Vol 3, Amsterdam: North-Holland, forthcoming.

[11] Fishman, Joshua (1977) "The Spread of English as a New Perspective for the Study of Language Maintenance and Language Shift" in Fishman, Cooper and Conrad (eds.) *The spread of English: the Sociology of English as an Additional Language.* Rowley, Mass.: Newbury House.

[12] Ginsburgh, Victor, Ignacio Ortuño-Ortín, and Shlomo Weber (2005) "Disenfranchisement In Linguistically Diverse Societies: The Case of The European Union", *Journal of the European Economic Association* 3, No. 4 (June), pp. 946-65.

[13] Grégoire, Henri (1794) *Sur la nécessité et les moyens d'annéantir les patois et d'universaliser l'usage de la langue française* in Michel de Certeau, Dominique Julia and Jacques Revel, 1975, *Une politique de la langue. La Révolution Française et les patois: l'enquête Grégoire.* Folio, Gallimard.

[14] Hagège, Claude (1996) *Le Français, histoire d'un combat.* Le livre de poche.

[15] Jackson, John H. (1938) *Finland*, George Allen and Unwin Ltd.

[16] John, Andrew and Kei-Mu Yi (2001) "Language and Location", Mimeo.

[17] Laitin, David D. (1994) "The Tower of Babel as a Coordination Game: Political Linguistics in Ghana", *American Political Science Review* 88, No. 3, pp. 622-34.

[18] Lang, Kevin (1986) "A Language Theory of Discrimination", *Quarterly Journal of Economics*, 101(2), pp. 363-82.

[19] Lazear, Edward P. (1999) "Culture and Language", *Journal of Political Economy* 107 (6), pp. 95-126.

[20] McRae, Kenneth D. (1983) *Conflict and Compromise in Multilingual Societies. Vol. 1, Switzerland.* Wilfried Laurier University Press.

[21] McRae, Kenneth D. (1986) *Conflict and Compromise in Multilingual Societies. Vol. 2, Belgium.* Wilfried Laurier University Press.

[22] McRae, Kenneth D. (1997) *Conflict and Compromise in Multilingual Societies. Vol. 3, Finland.* Wilfrid Laurier University Press.

[23] Mélitz, Jacques (2002) "Language and Foreign Trade" *CEPR Working Papers* No. 3590 (October).

[24] Ministère de l'Instruction Publique (1878) *Statistique de l'enseignement primaire. Tome second, Statistique comparée de l'enseignement primaire, 1829-1877.* Paris: Imprimerie Nationale.

[25] Pool, Jonathan (1991) "The Official Language Problem", *American Political Science Review* 85, No. 2, pp. 495-514.

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

[27] Kungliga Statistiska Centralbyrån (1924) *Statistisk årsbok för Sverige*, Stockholm.

[28] Statistiska Centralbyrån (2006) http://www.ssd.scb.se

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

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