IKERLANAK

THE ECONOMICS OF "WHY IS IT SO HARD TO SAVE A THREATENED LANGUAGE

by

Stefan Sperlich and José-Ramón Uriarte

2014

Working Paper Series: IL. 77/14

Departamento de Fundamentos del Análisis Económico I

Ekonomi Analisiaren Oinarriak I Saila

eman ta zabal zazu

UPV EHU

University of the Basque Country

The Economics of "Why is it so hard to save a threatened Language?"

Stefan Sperlich* José-Ramón Uriarte†

January 15, 2014

Abstract

We study the language choice behavior of bilingual speakers in modern societies, such as the Basque Country, Ireland and Wales. These countries have two official languages: A, spoken by all, and B, spoken by a minority. We think of the bilinguals in those societies as a population playing repeatedly a Bayesian game in which, they must choose strategically the language, A or B, that might be used in the interaction. The choice has to be made under imperfect information about the linguistic type of the interlocutors. We take the Nash equilibrium of the language use game as a model for real life language choice behavior. It is shown that the predictions made with this model fit very well the data about the actual use, contained in the censuses, of Basque, Irish and Welsh languages. Then the question posed by Fishman (2001), which appears in the title, is answered as follows: it is hard, mainly, because bilingual speakers have reached an equilibrium which is evolutionary stable. This means that to solve fast and in a reflex manner their frequent language coordination problem, bilinguals have developed linguistic conventions based chiefly on the strategy 'Use the same language as your interlocutor', which weakens the actual use of B.[1]
Keywords: economics of language, language conversation game, threatened languages.

1 Introduction

In a society with two official languages - denoted by A, spoken by all individuals, and B, typically spoken by a minority- it is said that there is a language contact situation. Lan-

*Département des sciences économiques and Research Center for Statistics, Université de Genève, Bd du Pont d'Arve 40, CH-1211 Genève, Suisse. E-mail address: `stefan.sperlich@unige.de`. The author acknowledges funding from the Swiss National Science Foundation 100018-140295.

†University of the Basque Country, Departamento de Fundamentos del Análisis Económico I, Avenida Lehendakari Aguirre 83, E-48015 Bilbao, Basque Country-Spain. E-mail address: `jr.uriarte@ehu.es`

[1]This joint work started while José-Ramón Uriarte was visiting Humbolt-Universität zu Berlin and Université de Genève. He would like to thank the generous hospitality of both universities. Financial support provided by the Basque Government and the Ministerio de Economía y Competitividad (ECO2012-31626) are gratefully acknowledged. Stefan Sperlich acknowledges financial support from the Swiss National Science Foundation, project 100018-140295.

guages compete for speakers, very much like firms compete for a market share. Language contact could be said to be the most extreme form of competition between languages. The pressure of the competition is felt, particularly, by the social support of language B, the minority of those who speak both official languages, the bilingual speakers[2]. The contact situation will shape the language choice behavior of this minority, the actual use they make of B in the interactions between them, their demand and supply of language B related goods and services, and the role they play in the transmission of B. Thus, the survival of language B and its related culture, and, therefore, the diversity of the society, depends solely on the bilinguals. Hence, what is at stake in this competitive situation is the society's language and cultural diversity. Quoting UNESCO's (2002) Article 1: 'As a source of exchange, innovation and creativity, cultural diversity is as necessary for humankind as biodiversity is for nature'.

The purpose of the present work is to give an answer to Fishman (2001)'s question, 'Why is it so hard to save a threatened language?', by means of the working tools available in economics, essentially, economic theory and econometrics.

The study of the dynamics of language competition was initiated by Abrams and Strogatz (2003). They showed that the long run outcome of two competing languages in a given population is that one of them will disappear. Their work gave rise to a substantial body of research, carried out, mainly, by physicists (an overview of this literature is Patriarca et al., 2012). Economists have been concerned, for the greatest part, with issues of language economics of a different nature; for instance, the relationship between earnings and language skills on markets where coexist several languages; see Chiswick and Miller (2007). An important part of those works, most of them empirical, are closely related to the Canadian-Quebec situation; see Vaillancourt (1980), Shapiro and Stelcner (1997) and Albouy (2008). On the theoretical side, economists have been attracted initially with the learning of a second language; the seminal work in this field is Selten and Pool (1991), which is the base of Church and King (1993), Ginsburg et al. (2007), Gabszewicz et al. (2011), that restricted the Selten and Pool's model to markets with two languages. They all have in common that the second language acquisition is modelled as a non cooperative game. On a different theoretical ground of language economics, we should point out Rubinstein (2000), who shows how the principles of communication efficiency shaping human language may give rise to certain binary relations that appear in natural languages; Blume and Board (2013) assumed that language competence is private information and that agents have different degrees of language competence giving rise to imperfectly shared meanings and uncertainty that affect the communication between them. They show that in games of 'common interest' it gave rise to severe efficiency losses. Chen (2013) has shown how the corpus of certain natural languages may shape the intertemporal economic behaviour of members of its speech community.

[2]In the present paper, a monolingual speaker does not become bilingual by learning any second language. It should be clear from the outset that we are referring only to bilingual speakers in the two 'internal' official languages A and B.

Grin et al. (2010) and Ginsburgh and Weber (2011) are two recent surveys of the most relevant research lines developed in the economics of language. We may conclude that there seems to be no work in language economics focusing on the question posed by Fishman (2001).

Fishman is referring to the fact that in a language contact situation the interactions between bilinguals are characterized by frequent word borrowings from the dominant language A, by constant language switching from B to A, and, more often, by a straight use of A. In other words, in a language contact situation, the social use of B, which is the key element for its survival, is less than what is statistically expected. One might think that this is due to, say, a lack of resources, poor education, or to bilinguals assigning a low status to B. To avoid simple explanations, we shall deal with quite the opposite case and set a benchmark. We will only consider threatened languages in a society satisfying the following two general features:

1. *The society is a democracy, highly developed economically.*
2. *The society has two official languages, A and B*[3]*, which are linguistically distant*[4].

Thus, in the societies satisfying these two conditions there are enough resources to design linguistic policies so that the decisions taken by the bilingual speakers might, to a certain extent, be implemented. This amounts to the existence of resources devoted to schools, teachers, textbooks, editing houses, media, institutions, that support the teaching and transmission of language B and its related culture, and markets where language related goods are traded. Further, B has become an official language because, mainly through voting, individuals reveal their linguistic preferences for B and claim their rights.

Thus the societies and languages satisfying the above conditions will set a kind of benchmark in the set of all societies with threatened languages contemplated in Fishman's question. Examples which would satisfy these features are the Basque Country, Ireland, Wales, and Scotland. In the Basque Country, the official languages are Basque and French in the French part, and Basque and Spanish in the Spanish part; in Ireland it is Irish and English; in Wales Welsh and English; in Scotland, Gaelic and English[5]. Of course, there are more examples satisfying the general assumption, but it is typically hard to get data which allow a deeper insight in the daily language use.

One would think that steady increases in the proportion of B speakers would imply similar steady increases in the social use of B. What happens in these officially bilingual societies is that increases in the knowledge of the minority language are accompanied by much smaller rates of increase in the use of B, and, in some cases, by an almost constant use.

[3] It is assumed too that B is official only in the concerned society.

[4] Therefore successful communication is only possible when the interaction takes place in one language.

[5] We do not include the case of French in Quebec because, 1. it is obvious that the fate of French and its related culture is not exclusively in the hands of the Francophones of Quebec. 2. French is a minority in overall Canada, but not inside Quebec and 3. within Quebec, some fractions of the Anglophones and Francophones are monolingual in their respective language.

We would say that in these societies, there seems to exist a kind of paradox, which we formulate as follows:

Why is it that having the political system, the legal instruments[6] to facilitate the use of B, the resources, and the education system to implement a language policy in favor of B, and - even more important - the people's support and willingness to speak the language, there is such a low use of B?

We propose the following framework to understand the issue and provide an answer to Fishman's question. First, note that, in the above mentioned countries, the population of bilingual speakers is a nice example of a real-life population playing a game that we call the Language Use Game (LUG). Bilingual speakers participate in frequent language choice situations, on a daily basis. In highly developed economies, interactions are, in most cases, anonymous, due to the mobility, both geographical and social, of the work force. This means that, often, bilinguals do not know ex-ante the linguistic type of the speech partners who are interacting with. External signals about linguistic types (monolingual or bilingual) in those societies are rare; for instance, accents signaling speakers of B are erased and all individuals in that society have essentially a similar accent shaped by the hegemonic language A. Thus, it is realistic to assume that linguistic types are private information. Bilinguals therefore face uncertainty as to which of the two languages they speak will be actually used in the interaction they are about to participate. In short, they must solve, under imperfect information, frequent language coordination problems and they have to activate the two linguistically distant languages (that is, recall equivalent concepts, names and meanings associated with the subject to deal with during the interaction) for an effective and efficient communication. Furthermore, bilinguals have the additional problem of maximizing their language preferences.

A bilingual would get the maximum payoff when he coordinates on his preferred language B, and would get the minimum payoff when he meets a monolingual speaker and is forced to use language A. The probability of the former event is given by the proportion α of bilingual speakers and the of latter is $1 - \alpha$. Hence, bilinguals' pure strategies are reduced to the following two: *Reveal the bilingual type by speaking B* - and, then you will participate in a lottery with the mentioned payoffs and probabilities - or *Hide the linguistic type* - and you will get, at least, the same payoff as a monolingual. The *Hide* strategy advices you to speak the same language as your interlocutor and, therefore, code switch from A to B only if the interlocutor happens to be a bilingual who plays *Reveal.* These are the main ingredients of the Bayesian game of language use played by the bilingual population. The game has a mixed strategy equilibrium, showing the optimal partition of the population of bilingual speakers into the fraction of those who play the *Reveal* strategy and that of those who play the *Hide* strategy. Members of the former group speak B when they interact with any other bilingual and those of the latter speak A between them.

[6] Examples are the Gaelic Language (in Scotland) Act of 2005; the Law of Normalization of Euskera's Use (in the Basque Country) of 1982; or the Welsh Language Measure of 2011, which gave the Welsh official status in Wales.

Based on this setting, we think of the equilibrium as a theoretical representation of the proportions of bilingual speakers who, in real-life situations, use language B in their interactions and those who do not. Then we build a function that relates to each proportion of bilinguals in $(0, 1)$ its corresponding Nash equilibrium in the game. We compare the predictions of language B use in equilibrium with the data about the measures of the actual use of three minority languages, Welsh, Irish and Basque. We find that the predictions obtained in this manner fit very well the empirical data about the (daily or street) use of the three languages.

Based on this, the answer we may give to Fishman is that a relevant part of the difficulties lie in that the bilingual speakers, as a player population, have reached an interior mixed strategy Nash equilibrium with strong stability properties: the equilibrium is evolutionary stable in the associated single population replicator dynamics. This means that bilinguals have built a linguistic convention, in which, typically, the group of those who play strategy *Hide* is larger than the group of those who play *Reveal*. The degree of dominance of the former group differ from country to country.

The rest of the paper is structured as follows. In Section 2, we present the concept of 'daily use' or 'street use' measure KE (making reference to the *Kale Erabilera* index defined for the Basque Country) for the minority language. Section 3 introduces the LUG -which is related to the Language Conversation Game of Iriberri and Uriarte (2012)- from which we elaborate a theoretical framework for the KE. In Section 4 we relate (or say, contrast) the theoretical analog of KE to a Nash equilibrium function that depends on the level of bilingual speakers. In Section 5 we estimate the model for the Basque Country, Ireland and Wales. We compare our empirical results based on our model with nonparametric analogues as a kind of model check, and study the results over time and countries. Section 6 concludes.

2 The 'Street Use' of a Minority Language

We are considering societies which have, essentially, two linguistic groups: the monolingual speakers, those who speak just language A and the minority of bilingual speakers, those who speak both official languages, A and B. In this setting, the minority language B is exposed to a direct competition with the majority language; that is, there is a language contact situation.

In this society, people, who are interacting at a certain time and place, could use either A, B, or even a mixture of both languages. Out of the total conversations that one could register at random, at a given time and place, one could count those conversations that used one of the two languages and know the proportion of people who took part in them. From there one may infer the proportion of the bilingual population of the observed place who use B in their interactions.

Definition of the Street Use Measure (KE):

Using random samples of anonymously registered conversations in the streets at a given time and place (say, a municipality or socio linguistic zone), the Street Use Measure (KE) of minority language B shows the number of individuals observed in conversations speaking language B out of the total number of individuals observed in the place. Dividing the KE index of the given municipality by the proportion α of bilingual speakers of that municipality, KE/α, we would obtain the Efficiency Index (EI).[7]

Thus, a municipality with a high EI means that the proportion of the bilingual population who actually uses language B is high. As this efficiency measure has already been used many times, we will keep this notation in this paper. However, we also introduce an additional index which we believe to be more informative and, as a byproduct, reveals a (in our case 'minor') problem with the KE index or any other 'daily use' index.[8] More specifically, we assume that the street use of language B refers to conversations of 'random matches', e.g. where the person starting a conversation does not have perfect information about the language abilities of the interlocutor. Therefore, an index that measures the 'efficiency of B' should relate the observed street use of B to the probability of randomly observing a conversation composed by two bilinguals. This probability is certainly Pr[bilingual]·Pr[bilingual]= α^2.[9] However, since this assumption of incomplete information and random matches is rather unlikely to hold in real life, we call this the 'hypothetical efficiency index' $HEI = KE/\alpha^2$.

Tables 1 to 3 show the α, KE, EI, and HEI for the case of the Basque[10] (on the Spanish side), Ireland[11], and Wales[12] for different years. Some features of the numbers are worth mentioning: The KE is explicitly observed only for the Basque Country; for Ireland and Wales these numbers are approximated by the recorded 'daily use'. We seem to observe a structural break for Ireland between 2002 and 2006. However, the simple truth is that only since 2006 we have information about the daily use of Irish outside the educational system. In other words, the KE is not really known for Irish before 2006. We should have in mind that in these tables only the aggregates are given; for the empirical study we will use the data taken on province or small area level.

[7]To our knowledge, the methodology for measuring the street use of a minority language based on anonymous observations has been developed by the group *Soziolinguistika Klusterra - the Sociolinguistic Cluster*, who operates in the Basque Country (see Altuna and Barturen, 2013)

[8]Note that in this paper 'daily use' refers to language use outside home, so that we again are in the 'street use' context. Certainly, the language use inside the educational system might then become a problem if it is not properly recorded, as we will see for the Irish data.

[9]Analogously, the probability of observing a conversation of two monolinguals is, under the above assumptions, $(1-\alpha)^2$, and the one of observing a mixture $2\alpha(1-\alpha)$.

[10]Data can be found on `http://www.soziolinguistika.org`. There is also given a detailed description of the definition and measuring of the *Kale Erabilera* index.

[11]Data are taken from `http://www.cso.ie/en/census/index.html`.

[12]All official statistics we found state only the language knowledge but no information about its use. The here used data are taken from Jones (2012) and counter checked with different reports released by the Welsh Language Board in Cardiff.

Table 1: Evolution of Knowledge α and Street Use (KE) of Basque in the Basque Country. First two columns give the percentage of bilinguals in the group of 16 years and over and the year this number was recorded by the Sociolinguistic Survey in the Basque Country. Columns three and four give the street use index (KE) and the year it was measured by the Cluster of Sociolinguistics.

Year	100α	Year	$100KE$	EI	HEI
1991	22.30	1993	11.80	0.53	2.37
1996	24.40	1997	13.00	0.53	2.18
2001	25.40	2001	13.30	0.52	2.06
2006	25.70	2006	13.70	0.53	2.07
2011	27.00	2011	13.30	0.49	1.82

Table 2: Evolution of Knowledge α and Street Use (KE) of Irish in the Republic of Ireland. First two columns give the percentage of bilinguals in the group of 3 years and over and the year this number was recorded by the different Census in Ireland. Columns three and four give the street use index (KE) and the year measured.

Year	100α	Year	$100KE$	EI	HEI
1996	41.10	1996	10.16	0.25	0.60
2002	41.88	2002	09.05	0.22	0.52
2006	40.83	2006	02.10	0.05	0.13
2011	40.60	2011	02.15	0.05	0.13

Note that where we have observations over time - 1991 to 2011 for the Basque Country, 1996 and 2002 for Ireland, and 2006 and 2011 for Ireland - we see that the efficiency index EI is quite stable, except for 2011 in the Basque Country where it lost 4 percentage points. When looking at the hypothetical efficiency index HEI, we observe superefficiency, i.e. values above 100%. In other words we observe by far larger proportions of bilingual conversations than one would expect if all conversations were random matches with incomplete language information. The only possible explanation is that this assump-

Table 3: Knowledge α and Street Use (KE) of the Welsh in Wales. First two columns give the percentage of bilinguals in the group of 3 years and over and the year this number was recorded by the Population Survey in Wales. Columns three and four give the street use index (KE) and the year it was estimated from the Language use surveys.

Year	100α	Year	$100KE$	EI	HEI
2005	26.60	2005	15.38	0.58	2.17

tion might fail what would explain these high values for HEI. What is more likely and sufficient for our further analysis is, that a certain but fixed percentage, say $(1-p)$, of all conversations is not such a random match. Having this in mind, we can conclude from the tables that the efficiency index of interest should, HEI, i.e. the one accounting not just for the proportion of bilinguals but for the proportion of conversations composed of two bilinguals, has been steadily decreasing. This is especially surprising for the Basque Country where the percentage of bilinguals has had a big increase (from 1993 to 2011 by 5 percentage points, that is, by more than 20%).

Any theoretical model intended to have a certain descriptive power for the minority language street use should capture all these findings and aspects of the KE, EI and HEI corresponding to the language under investigation. In the next section we propose a mathematical representation for the 'minority language street use'.

3 A Model for the Street Use of a Minority Language.

In the context of a society with a language contact situation, bilinguals must make frequent language choices and are, in a natural way, involved in a game of language coordination, seeking to maximize their language preferences and communication efficiency. In this section we show how the bilingual speakers could be thought of as a population playing a game of language use. From there we derive the main hypothesis of the paper.

The main features of the society we are dealing with are captured in the following assumption.

General assumption: *Economically, the society is highly developed, and politically, it is a democracy. Linguistically, the society has two official languages, A and B, which are very distant, so that successful communication between members of the society is only possible in one language.*

A feature of highly developed modern societies is the mobility, both social and geographical, of the work force. This gives rise to frequent anonymous interactions; thus, we will assume:

Imperfect information: *in a given percentage p of conversations, the participants of an interaction do not have, ex-ante, any information about the linguistic type (bilingual or monolingual) of each other. They only know the proportion of bilingual and monolingual speakers, α and $(1-\alpha)$ respectively, of the society.*

Democracy allows, mainly through voting, to know the individual preferences which are the basis for collective decisions. The fact that language B is official reflects not only the linguistic rights of the minority language speakers, but, also, their preference (weak or strong) for language B. Thus, more formally, we assume:

Language loyalty or preference: *bilingual speakers prefer to speak B.*

On a first glimpse this assumption might be disputable to a certain extent. However, it will automatically be relaxed by a proper choice of the payoffs which reflect how strong or weak this preference really is (we develop this issue in the next section).

A monolingual speaker does not make language choices and, thus will always get a sure payoff, say, n. Since choices are made under imperfect information, a bilingual may choose the majority language A; in that case, we will assume that she will get, as a monolingual, the payoff n, because this was a voluntary choice. Bilingual speakers will get the maximum payoff, m, when they coordinate in their preferred language B. However, $(n-c) > 0$ would be the payoff to a bilingual speaker who, having chosen B, is matched to a monolingual and, therefore, is forced to speak A; then c denotes the frustration cost felt by this bilingual. We make the following assumption about payoffs' ordering:

Payoffs: *For a given proportion $\alpha < 1 - \alpha$ we assume $m > n > c > 0$. Further, the frustration cost should be smaller than the weighted benefits, i.e. $c < (m-n)\frac{\alpha}{(1-\alpha)} =: b(\alpha)$.*

Under this set of assumptions, the bilinguals' language behavior is captured fairly well by the following strategies:

s_1: *Use always B, whether you know for certain you are speaking to a bilingual individual or not. Use A only when the speech partner reveals he is of the monolingual type.*

s_2: *Use B only when you know for certain that you are speaking to a bilingual individual; use A otherwise.*

Notice that playing s_1 the bilingual reveals his type, whereas playing s_2 the type remains hidden. With s_1 you risk to get the minimum payoff, $n - c$; bu s_2 is not risky because it advices you to "speak the same language as your interlocutor"; that is, use A and switch to B only if the other party plays s_1. Thus, with s_2 you expect to get at least n.

The Bayesian game in which bilinguals are involved could be explained as follows (see Figures I and II). A bilingual expects to meet another bilingual with probability α and a monolingual with probability $1 - \alpha$. In the former event, each bilingual will play one strategy, s_i $(i = 1, 2)$, get a payoff, and speak a language, A or B; clearly, in this event s_1 weakly dominates s_2. In the latter event, the bilingual will get a payoff depending on the chosen strategy and will speak, in any case, language A; now strategy s_2 strictly dominates s_1.

Let us now interpret the game as one player population game: a game played by the population, N, of bilingual speakers. Let x be the proportion of bilingual speakers who play strategy s_1. Under the assumptions of Imperfect information, Language loyalty and Payoff ordering the Bayesian game illustrated also in Figures 1 and 2 has a mixed strategy Nash equilibrium[13] in which the bilingual population plays s_1 with probability

$$x^* = 1 - \frac{c(1-\alpha)}{\alpha(m-n)} \qquad (1)$$

[13] It can be seen that the expected payoff matrix associated to the game has two additional equilibria, (s_1, s_2) and (s_2, s_1), which are unstable.

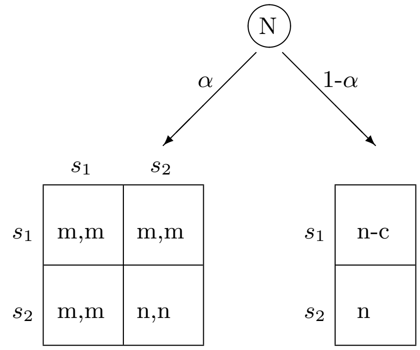

Figure 1: The Language Use Game.

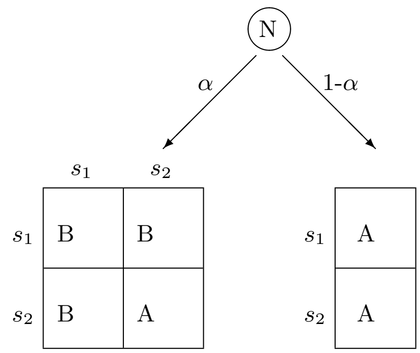

Figure 2: The languages spoken in the Language Use Game.

Iriberri and Uriarte (2012) show that x^* is evolutionary stable in the associated one-population Replicator Dynamics. Then the equilibrium could be thought of as an optimal partition of $N = Nx^* \cup N(1-x^*)$ and the partition could be considered as a 'linguistic convention' built, in the long run, by the bilingual speakers; see Weibull (1995). Indeed, the language strategy played by each group in equilibrium is the following:

1. The subpopulation Nx^* consists of bilingual speakers who reveal their bilingual type by playing the pure strategy s_1. Members of this group will speak B when they interact with other bilinguals.

2. The subpopulation $N(1-x^*)$ consists of those who 'hide' their bilingual type by playing the pure strategy s_2.

Hence, in the interactions between bilingual speakers belonging to this group the language spoken is A. Members of this group will only speak B when they interact with those in Nx^*.

In Tables 1 to 3 it is distinguished between those who know the minority language (that is, the proportion α of bilingual speakers) and those who actually use it (that is, the KE measure). The distinction between *knowledge* and *use* of language B is well captured by the present game. The equilibrium of the game is telling us the number or percentage of bilingual speakers who actually use B when they interact, i.e. those in group Nx^*. This leads us to the following

Hypothesis:

The mixed strategy Nash equilibrium x^* could be thought of as a theoretical representation (that is, a model) of the fraction of bilingual speakers who, in real-life situations, use language B in their interactions.

Actually, we know the precise number of those who have the knowledge of the minority language B, and we have observations of KE. Hence, in the next section, we shall model x^* as a function depending on the proportion of bilingual speakers, $\alpha \in (0, 1)$, and take KE as an observation for $\frac{\alpha}{p}x^*$ that approximates the actual use of B (where p is the proportion of conversations under imperfect information). Note that x^* is the proportion of bilingual speakers who play strategy s_1, whereas $p \cdot EI = KE \cdot p/\alpha$ is its empirical counterpart in the equilibrium.

Thus, the predictions one might obtain with our model for x^* should fit well the data about the actual street use of B for all the α and KE values we observe on the province and/or small area levels in the different sociolinguistic zones of the considered societies (namely the Spanish Basque Country, the Republic of Ireland, and Wales). The purpose now is first to establish a model and then to confront this hypothesis with the empirical evidence.

4 The Street Use Measure as an Endogenous Nash Equilibrium function

The parameters m, n and c in (1) are exogenously given, but it seems natural to assume that as α increases, both the payoff m bilingual speakers get when they interact in their preferred language B, and the frustration cost c, should rather decrease than increase. That is, when α reaches a certain higher level, B would then be perceived, mainly by its speech community, not as an endangered language but as a normalized one. As α keeps increasing, bilingual speakers would tend to feel that there are no reasons for exceptional levels of utility or payoffs and would be inclined to assign smaller payoffs to the (now a much more frequent) event of coordinating in language B. In other words, as α increases, m would decrease and approach the payoff level, n, of the normalized language A. By the same reason, the event of failing to coordinate in B would be less frequent, and so the frustration cost c would decrease too and tend to 0. Hence, there must exist some crucial level for alpha, which we denote as α^* at which the convergence of m to n and of c to 0 will occur.

One would expect that $x^* = x^*(\alpha)$ would be an increasing function of α. That is, as the proportion of bilingual people increases, the equilibrium proportion of bilingual speakers who play strategy s_1 will also increase.

The payoff n is obtained by the monolingual speakers and by those bilingual speakers who voluntarily chose to speak A. We shall consider n as the natural payoff level that one might obtain from using a socially normalized language, such as A, the language spoken by all members of the society. For this reason, we shall keep the payoff n constant.

Presumably, the model of minority language use would gain more descriptive power if both m and c were decreasing functions of α. Therefore, one should choose functional forms being decreasing in α and compatible with $x^* = x^*(\alpha)$ be an increasing function.

Given certain functional forms of $m(\alpha)$ and a value for n, the weighted profit $b(\alpha) = (m(\alpha)-n)\frac{\alpha}{(1-\alpha)}$ would be decreasing, even though the ratio $\frac{\alpha}{(1-\alpha)}$ is an increasing function of α. Let us suppose that $m(\alpha)$ has the simple functional form $m(\alpha) = \frac{K}{\alpha}$, where $K > 0$ is a constant.

Let α^* denote the proportion at which bilingual speakers perceive that B has reached the status of a socially normalized language in the sense that $m(\alpha^*) = n$. For any $\alpha < \alpha^*$, $m(\alpha) > n$, and a bilingual speaker would get a positive net profit $m(\alpha) - n$ whenever he is able to coordinate in language B.

As it was said above, the frustration cost $c(\alpha)$ should be a decreasing function of α; this would be the case if also the weighted benefit function $b(\alpha)$ is decreasing. Furthermore, to allow the use of B in the equilibrium, we assumed $0 < c(\alpha) < b(\alpha)$ for any $\alpha < \alpha^*$. We get

$$c(\alpha) = (m(\alpha) - n)\frac{\alpha}{(1-\alpha)} - R(\alpha) \ , \tag{2}$$

with $R(\alpha) > 0$ being the net benefit. Now, inserting equation (2) in (1) we obtain

$$x^*(\alpha) \quad = \quad \frac{\alpha(m(\alpha) - n) - c(\alpha)(1-\alpha)}{\alpha(m(\alpha) - n)} = \frac{R(\alpha)(1-\alpha)}{\alpha(m(\alpha) - n)} \ .$$

And with $m(\alpha) = \frac{K}{\alpha}$,

$$x^*(\alpha) = \frac{R(\alpha)(1-\alpha)}{K - n\alpha} \ . \tag{3}$$

Equation (3) shows how the equilibrium proportion of the bilingual population playing s_1 changes as the proportion of bilingual speakers, α, changes. Note that the denominator of (3) is greater than zero because $m(\alpha) = \frac{K}{\alpha} > n$ for all $\alpha < \alpha^*$.

Hence the function $x^* = x^*(\alpha) > 0$, will be increasing in α if $R(\alpha)$ is either increasing or at least not decreasing faster than $\frac{1-\alpha}{\alpha(m(\alpha)-n)}$ is increasing in α.[14] Therefore one would suppose that the net benefit function

$$R(\alpha) = b(\alpha) - c(\alpha) = (m(\alpha) - n)\frac{\alpha}{(1-\alpha)} - c(\alpha)$$

is a function of α with a well defined first derivative. The perception that B has reached the status of a normalized language will normally occur before a 100% of the population becomes bilingual, making $m(\alpha^*) - n = 0$, at some $\alpha^* < 1$.

However, people may have different perceptions about when the minority language B could be said to be normalized. Given a certain linguistic context (say, a certain municipality

[14]It can easily be checked that $g(\alpha) = \frac{1-\alpha}{\alpha(m(\alpha)-n)}$ is increasing by calculating its first derivative, inserting $m(\alpha) = K/\alpha$ and using that $K/\alpha^* = n$ with $\alpha^* < 1$.

or linguistic zone), the perception of an individual would be conditioned by that language environment. The present proportion of bilingual speakers in that environment will define the minority language reference point for each individual living in that context. Then people will perceive and quantify α^* depending of that reference point. As it happens with the perception of attributes such as wealth, the frequency with which an individual experiences the event of meeting bilingual speakers in the past and in the present will determine an adaptation level or reference point. Quoting Kahneman and Tversky (1979): "The same level of wealth, for example, may imply abject poverty for one person and great riches for another, depending on their current assets". Similarly, in a linguistic context where the number of bilingual speakers is relatively small, people would assign a value to α^* smaller than the value assigned by those who live in a context with a relatively high proportion of bilingual speakers.

Let $\alpha_N(\alpha)$ be the function that assigns to each α of a linguistic context N the value α^* at which the B speakers of that area perceive that the language B is already normalized. We shall assume that $\alpha_N(\alpha)$ is a concave and increasing function with negative second derivative, as the 'value function' of Kahneman and Tversky (1979). Thus, $\alpha_N(\alpha) = \alpha^*$, with $\alpha^* > \alpha$, would be the value at which $m(\alpha^*) = n$. Each linguistic context N will have a specific α^* or perception as to which is the proportion of language B speakers that converts B into a normalized language.

Then, at the convergence point, $m(\alpha^*) = \frac{K}{\alpha^*} = n$, so the constant $K = n\alpha^*$. Making the substitution in the denominator of (3), we get:

$$x^*(\alpha) = \frac{R(\alpha)(1-\alpha)}{n(\alpha^* - \alpha)} \ . \tag{4}$$

Note that since $x^*(\alpha)$ is an interior mixed strategy equilibrium, that is, $x^*(\alpha) \in (0, 1)$, then $R(\alpha)(1-\alpha) < n(\alpha^* - \alpha)$, where $1 > \alpha^* > \alpha$, and hence $n > R(\alpha)$.

Corollary

Let α denote the proportion of bilingual speakers in a certain sociolinguistic context. Then $\frac{\alpha}{p}x^*(\alpha) = PKE(\alpha)$ is the predicted street use of B (PKE) in that sociolinguistic context.

Putting together all calculations and considerations from above, this function $PKE(\alpha)$ is

$$PKE(\alpha) = \frac{\alpha R(\alpha)(1-\alpha)}{pn(\alpha^* - \alpha)}$$

With specifications $R(\alpha) = \beta_1 \alpha^{\beta_2}$ for unknown β_1, β_2, and $\alpha^* = \alpha^{b_3}$ (b_3 also unknown), we get

$$PKE(\alpha) = \frac{\alpha \beta_1 \alpha^{\beta_2}(1-\alpha)}{pn(\alpha^{b_3} - \alpha)}$$

which for unknown n, p is obviously not identified. In fact, without further information from outside the model, it is only identified up to

$$PKE(\alpha) = \frac{b_1 \alpha^{b_2}(1-\alpha)}{(\alpha^{b_3} - \alpha)} \tag{5}$$

with $b_1 = \beta_1/(np)$ and $b_2 = 1 + \beta_2$. This is the model we shall study empirically, based on data for the Basque, Irish and Welsh. Note that there is no particular reason why n, the pay-out for communicating in A, or p, indicating the percentage of random matches among all conversations, should have seriously changed over the considered time window. Consequently, even tough β_1 is not identifiable, you can perfectly interpret the development of b_1 over time as the development of β_1. Moreover, you might even assume that np does not vary substantially over the considered regions (Wales, Ireland and Basque Country); in that case you can also interpret the differences in b_1 between regions as the differences in β_1. In other words, we can identify the differences and changes in the net benefit function $R(\alpha)$ over time and region once we assume np to be constant.

5 Empirical Evidence

For each of the three languages and regions we will present in the following (a) the estimated parameters of function (5), (b) the resulting functional forms of PKE, compared with a nonparametric fit of the observed KE on α. We will further study the distributions of α, KE and EI over the provinces or/and small areas for each considered language.

In order to estimate function $PKE(\alpha)$ from the samples $\{KE_{cti}, \alpha_{cti}\}_{i=1}^{n_{ct}}$ for country sample c in year t one might consider either

$$KE_{cti} = \frac{b_{ct,1}\alpha_{cti}^{b_{ct,2}}(1-\alpha_{cti})}{(\alpha_{cti}^{b_{ct,3}} - \alpha_{cti})} + \varepsilon_{cti} \tag{6}$$

or, as it could be thought, as a model with multiplicative structure,

$$\log(KE_{cti}) = \log(b_{ct,1}) + b_{ct,2}\log(\alpha_{cti}) + \log(1-\alpha_{cti}) - \log(\alpha_{cti}^{b_{ct,3}} - \alpha_{cti}) + \epsilon_{cti}$$

and estimate the parameters $b_{ct,j},\ j = 1, 2, 3$ by non-linear least squares under the constraints that $0 \le b_1 < 1$ and $b_3 < 1$.[15]

While the general findings are quite similar for one or the other estimation strategy, predicting the KE from the logarithmic version (and consequently $\widehat{\log(KE)}$) is somewhat more complex as one has to correct for the - in our case heteroscedastic - error dispersion since $E[\log KE|\alpha] < \log E[KE|\alpha]$. We therefore concentrate on the presentation of the least square estimates resulting from model (6)[16]. All figures are given together with nonparametric fits of KE on α using local quadratic estimators with Epanechnikov kernel and local bandwidth such that 25% of all sample points are inside the kernel support.[17] For details see the Appendix.

[15] From our discussion above, we see that there is no clear constraint for b_2, although one would expect that $b_2 - 1 = \beta_2 > 0$, giving an increasing $R(\alpha)$.

[16] The other results are available on request from the first author.

[17] More specifically, we used the **R**-procedure `locfit` with $\alpha = 0.25$ and $deg = 2$.

5.1 Estimation Results

For the estimation of the PKE function we considered it as inadequate to weight the observations made by provinces, municipals or small areas by their population size. The reasons are manifold, and to discuss them is beyond the scope of this exercise; we let it with the remark that for studying the theoretical model each combination of (α, KE) is for us an equally valid information. We are less interested in the parameter estimates for a country or language than in the question of how well the language game model explains.

We start with Wales for which we have reliable data only for 2005.[18] In Figure 3 and Table 4 are given the results of estimating the PKE, respectively equation (6), for Wales. The solid line refers to the parametric model, the dashed one to the nonparametric analogue. The circles indicate the recorded observations.

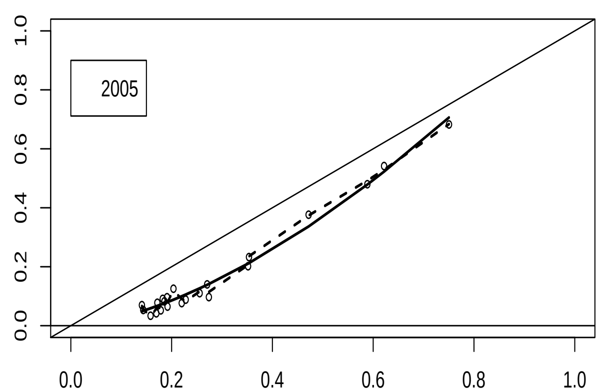

Figure 3: Parametric (solid line) and nonparametric (dashed line) estimates of PKE for Wales, together with angle bisector.

Table 4: Parameter estimates of equation (6) for Welsh local authorities; *obs.* indicates the number of observations.

year	2005
b_1	1.078
b_2	1.618
b_3	0.035
obs.	22

Certainly, given the small sample size, the variance of the nonparametric estimator is expected to be pretty large. Nonetheless we see and can conclude from the main characteristics: there is no major difference between the parametric and the nonparametric curvatures which both are well adapted to the data. That is, our model for x^* seems to fit pretty well what has been observed regarding the street use of the Welsh language. Given

[18]The data were actually collected during the period from 2004 to 2006.

the fact that we have no information about p and n (percentage of random matches and payoff of conversations in English), there is no particular interpretation for b_1 but we know that $\alpha^* = \alpha^{b_3}$ while $\beta_2 = b_2 - 1$ gives us the speed at which the net benefit function $R(\alpha)$ increases with α. The net benefit increases at a faster rate than $\sqrt{\alpha}$ but with decreasing intensity ($\beta_2 < 1$). The percentage α^* at which Welsh is no longer perceived as a minority language (such that $m = n$) seems to be above 90%.

For Ireland, we have data for different levels of aggregation, namely for about 3400 so called 'electoral divisions', for the about 180 'local electoral areas', and for the 34 counties. The last aggregation level is of little help as it exhibits little variation in α. The estimation results are given in Table 5 and Figure 4.

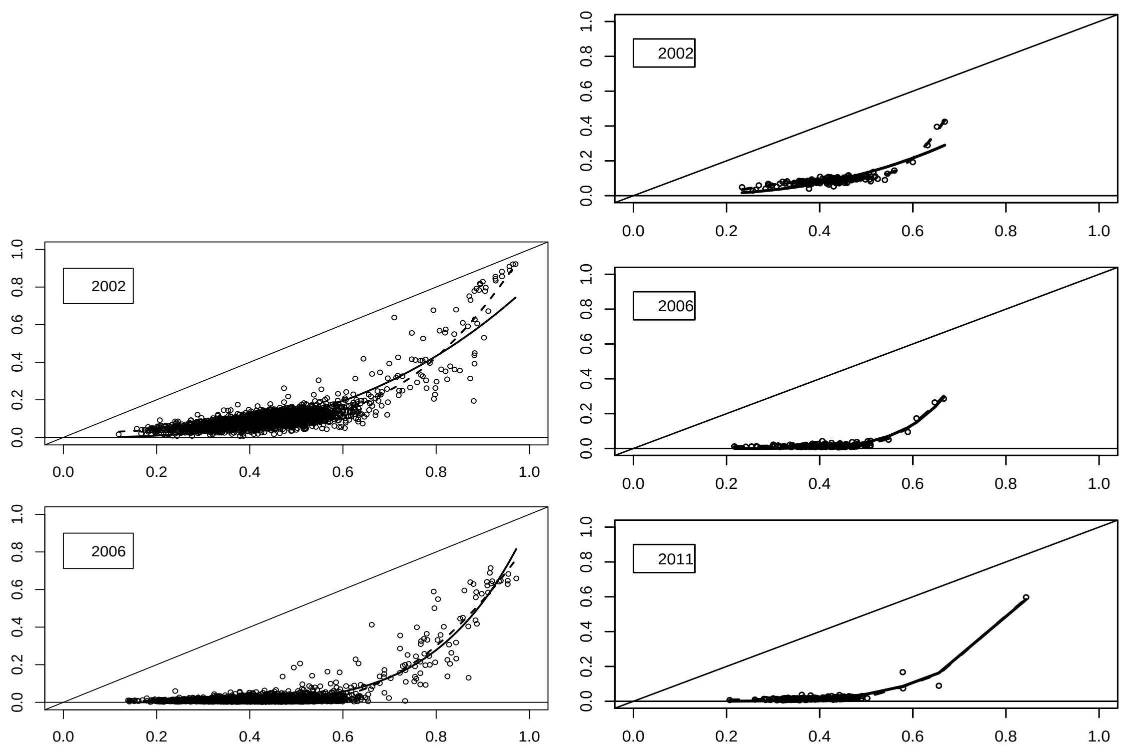

Figure 4: Parametric (solid line) and nonparametric (dashed line) estimates of PKE for Ireland, together with angle bisector: right panel for 'electoral divisions', left panel for 'local electoral areas'.

Admittedly, for both the Irish and the Welsh data there are some uncertainties concerning the KE. A main problem is that daily use is not necessarily a good measure for KE. As already discussed for the Irish data, people may use every day language B only at school but elsewhere play strategy s_2. Only since 2006 there is a clear definition of daily use outside the educational system. Consequently, the difference between the parametric and the nonparametric fit for Ireland in 2002 might be simply to a miss-measurement of KE. We see, however, that for 2006 and 2011 our model fits pretty well the observed data of daily language use being close to the nonparametric estimate (data fit without a model).

Concerning the parameter estimates we notice that different aggregation levels lead to quite

Table 5: Parameter estimates of equation (6) for Irish electoral divisions (first two columns), and Irish local electoral areas (last three columns); *obs.* indicates the number of observations available for that year.

year	2002	2006	2002	2006	2011
b_1	0.029	0.329	0.202	1.909	0.644
b_2	3.300	5.863	3.149	7.684	5.381
b_3	0.964	0.655	0.769	0.669	0.538
obs.	3422	3409	180	180	201

different estimates for b_1 what is not surprising as these may lead to differently perceived payoffs (n), and maybe also to different likelihoods of random matches (p). Fortunately, the two parameters with some value of interpretability, b_2 and b_3, are comparable over the different aggregation levels. We conclude that the net benefit function $R(\alpha)$ has become much steeper with regard to α (from about $\beta_{1,2002}\alpha^2$ in 2002 to about $\beta_{1,2006}\alpha^5$ in 2006), but as $\alpha < 1$, the perceived net benefits have actually diminished a lot. The α^*, where $m = n$, went down from almost α to just a bit more than $\sqrt{\alpha}$.

For the Basque country we have data where a correct measurement of KE is guarantied. Moreover, the KE measure did not change so that we can airily compare the five years to study the long-term dynamics (over a period of almost 20 years) later on. The estimates are shown in Figure 5 and Table 6. In the Figures, the non-parametric estimates are shown as dashed lines whereas the solid lines are the $x^*(\alpha)$ model with the parameter estimates as shown in Table 6.

Table 6: Parameter estimates of equation (6) for the Basque municipals; *obs.* indicates the number of observations available for that year.

year	1993	1997	2001	2006	2011
b_1	0.136	0.116	0.118	0.176	0.685
b_2	2.192	1.894	2.204	2.262	2.559
b_3	0.838	0.858	0.864	0.815	0.434
obs.	101	121	134	74	84

A main difference with respect to the other linguistic zones above, is that the dispersion is much larger here: for α as well as for the α conditioned KE. This gives the misleading impression that our theoretical model would capture less well the reality. As long as we take α as the only varying explanatory variable, this is not true: the nonparametric fit and the theoretical model are very close; in fact, the nonparametric one is just a bit more wiggly. This indicates that it is not possible a better projection of KE on α than the one we provide with our theoretical model for x^*. However, the additional, but unexplained

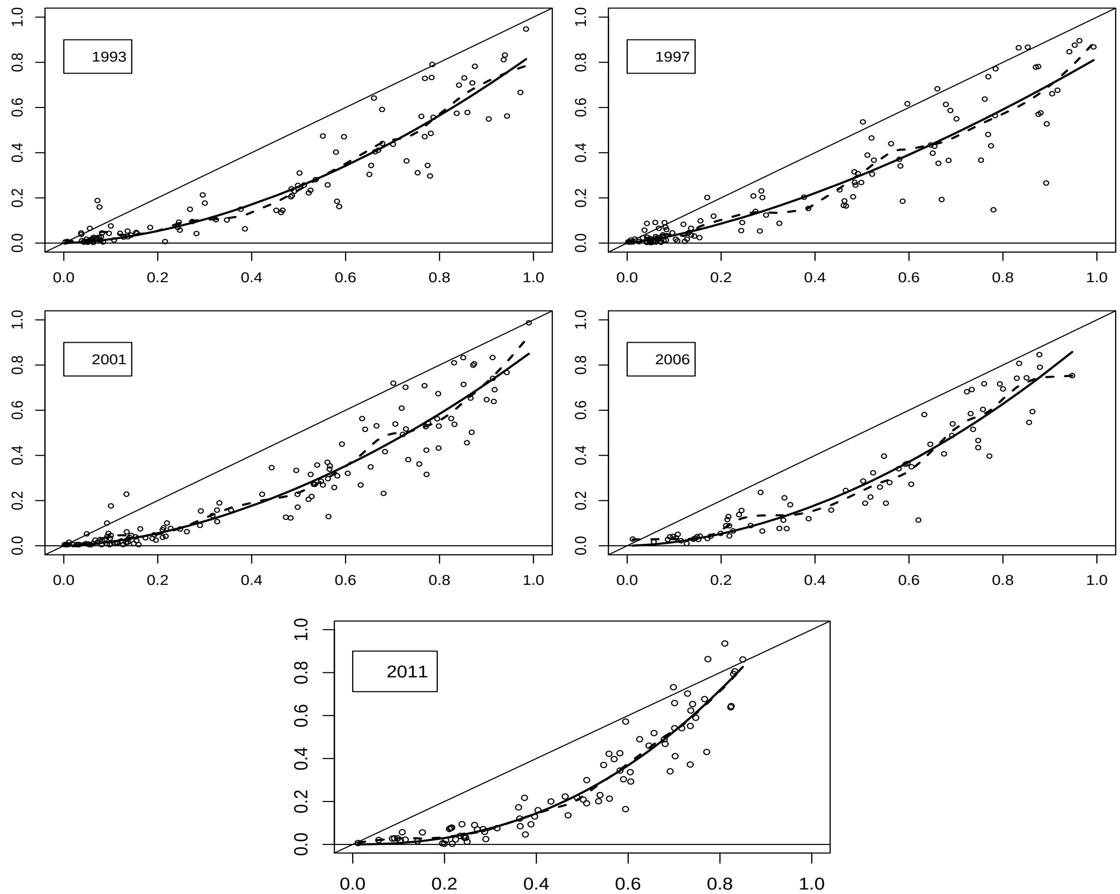

Figure 5: Parametric (solid line) and nonparametric (dashed line) estimates of PKE for the Basque country, together with angle bisector.

variation in KE for given α, may easily be explained by local particularities. To look for those and include them in a regression model might be an interesting exercise, but it is not the aim of our language game model.

When looking at the parameters, we see a main change only from 2006 to 2011, but only for b_1 and b_3. That is, while we hardly see a change in the net benefit function $R(\cdot)$, the α^* went down from about $\alpha^{0.85}$ to $\alpha^{0.43}$. This is a similar development as we observed for Ireland in what concerns α^*, but is contrasted by the stable net benefit $R(\cdot)$, which actually has even increased for the Basque Country due to the in average increasing α. Recall that in Ireland this net benefit went down to almost zero, what strongly fosters the extinction of the Irish language. For the Basque we only have a slight, probably insignificant increase in b_2 giving a $\beta_2 \approx 1.6$ what results in an increasing $R(\alpha)$ with increasing returns to α. Again, for b_1 we have no clear interpretation - and so we have for its change in 2011.

5.2 Tests and Analyzing the Changes over Time

We also tested the functional form of our model nonparametrically for each country and year. This was done along the bootstrap test of Härdle and Mammen (1993), see appendix. One has to know that nonparametric tests conditioned on the design - like almost all nonparametric bootstrap tests are - will always reject once the sample size is large enough compared to the residuals variance. It is then up to the empirical researcher to decide whether the detected statistically significant differences matter for his research question or not. In our case we can see that for example for the Irish data of 2002, i.e. when KE is by far over-reported, our model can not replicate this shift as it does not contain an intercept. Consequently, for the Irish data of 2002 the test is expected to reject. For 2006, the Irish data exhibit a slightly stronger bend (like of an elbow) than the parametric model can produce so that given the astonishingly small residual variance and given the sample sizes, the test should (at least 'almost') reject. For all other years and data we expect to not reject at a 10% or even 20% level. These were actually exactly the results we obtained when performing the test based on 100 random wild bootstrap samples.

As for the Basque Country we have the most reliable data with data available over almost two decades, we can also study the dynamics over time. We already discussed the development of parameters b_1, b_2, b_3. In Figure 6 we have summarized the changes of the PKE function over time. Though quite stable over the years, we mainly see that the street use of Basque for given α seems to steadily increase for municipals where $\alpha > 0.5$ whereas for those with $\alpha < 0.5$ it is varying over time without a clear tendency. This finding we make independently from looking at our model or the nonparametric estimates.

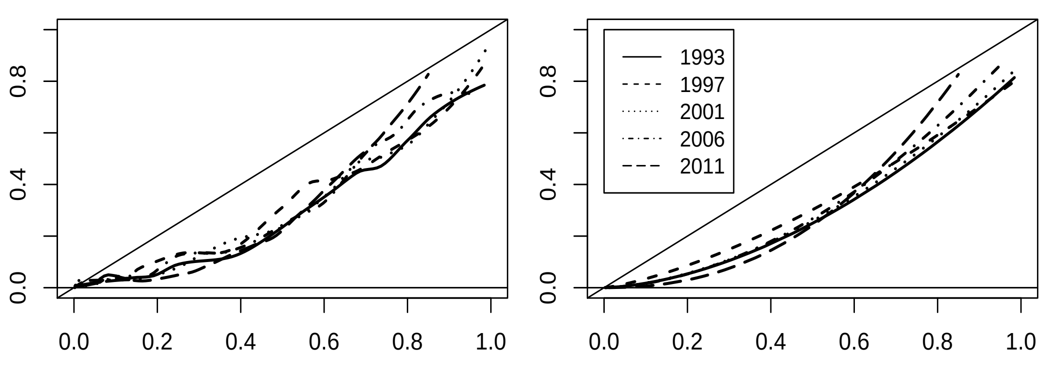

Figure 6: Nonparametric (left panel) and parametric (right panel) estimates of PKE for the Basque country.

Having said this it would be interesting to contrast this with the development of the α but also the KE and EI, each separately. The box-plots in Figure 7 illustrate quite well the development of the distributions over the years. First, recall that we are looking at all combinations (α_{ti}, KE_{ti}) (for $t = 1993, 1997, 2001, 2006, 2011$) without weighting them by the population size of municipality i. This explains why it seems that the percentage(s) of bilingual speakers went down though the real total percentage has steadily increased, see

Table 1. We see that all years exhibit a huge dispersion for α and KE with no stabilization of any of the distribution of these indices. We observe a shrinking number of municipalities with only small α or/and small values of KE.This might explain why people now feel that α has still to increase quite a bit (i.e. b_3 has fallen) to become a normalized language (i.e. $\alpha = \alpha^*$). At the same time, the slightly increased net benefit in 2011 is not clearly reflected in these box-plots.

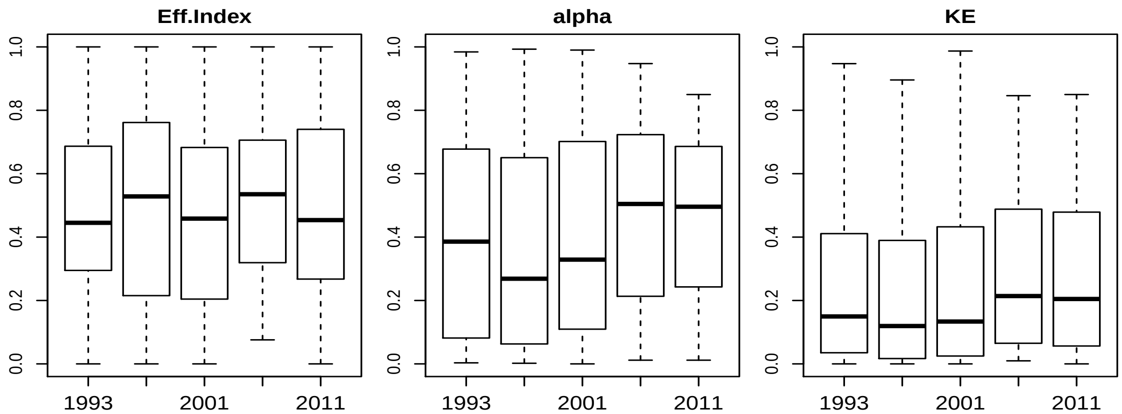

Figure 7: Boxplots of the distributions of the indices α, KE, and $EI = KE/\alpha$ over the regions for each observation year in the Basque Country.

The same analyzes for Ireland and Wales show simply in a different way what we already found looking at the tables and figures above, and are therefore skipped.

6 Conclusions

We have built the $x^*(\alpha)$ function that relates the proportion of bilingual speakers, $\alpha \in (0,1)$, with the (Nash) equilibrium proportion of bilingual speakers, x^*, who play the strategy s_1 : *"use always language B ..."* in the Bayesian Language Use Game. We think of this function as a model for the street use measure, *KE,* of the minority language *B*. We show that the predicted street use of language B, $PKE(\alpha)$, is a strictly increasing convex function on α. That is, $x^*(\alpha)$ captures the empirical fact that the use of B increases with α. Thus, our model predicts an equilibrium use of language B with strong stability properties; that is, $x^*(\alpha)$ is evolutionary stable, as well as asymptotic stable in the associated one-population replicator dynamics.

When we study the data about the actual street use of Welsh, Irish and Basque languages, we observe a relationship between percentage of bilingual speakers α and street use which is as predicted by the theoretical model. Moreover, while the parameters change considerably, when comparing the PKE forms and locations they are pretty stable over the years, though not over the countries. However, the latter might be attributed to the different ways of having measured the street use of the minority language in question.

The parametric model has been compared to nonparametric (model-free) fits of the observed KE on their corresponding α. The functions resulting from the theoretical model came astonishingly close to these model-free data fits. This holds also true over time. As we indicated above, since the equilibrium use of language B is evolutionary stable, it could be interpreted as if the bilinguals build linguistic conventions to solve their language coordination problem under imperfect information.

Hence, given the specifications of the model, to the question posed by Fishman (2001) "Why is it so hard to save a threatened language", we would say that it is mainly because bilinguals face frequent language choice decisions to coordinate language with interlocutors of unknown linguistic type. Thus, they are in the need of decision procedures to solve fast that coordination problem. Then, by interactive learning, bilinguals reach an evolutionary stable equilibrium or, equivalently, a linguistic convention which, typically, is strongly based on the strategy "hide your linguistic type". We show that this strategy reduces the use of B.

The linguistic convention introduces a strong stability component into the linguistic behavior of the bilingual population that is hard to break. Roughly speaking, it would be needed political measures to either increase the bilinguals' perceived net benefit of using B or to reduce the imperfect information. A dramatic increase in the proportion α of bilingual speakers is obviously not the key point, as has been proved by comparing Ireland, Wales, and the Basque Country: while the percentage of bilinguals in Ireland doubles the one in the Basque Country, the former is close to extinction while the latter exhibits a pretty stable street use.[19] One might speculate that this is because English is the competitor, i.e. the majority language A, is much more dominant, what makes it particularly hard for the Irish to survive.[20] For this reason we added the Welsh; it has a comparable α like the Basque but even a slightly higher KE (what might be simply due to the different measurement). Unfortunately the aggregation level for Wales is too high to draw many conclusions from the model parameter estimates.

References

Abrams D.M. and Strogatz S.H. (2003) "Modeling the dynamics of language death", *Nature* 424: 900 pp.

Albouy, D. (2008) "The Wage Gap Between Francophones and Anglophones: a Canadian Perspective, 1970-2000", *Canadian Journal of Economics*, vol. 41, no. 4, pp. 1211-38.

[19]Some people might expect it to increase more along α, but recall that the probability of a bilingual random match is α^2, not just α. In that sense it is correct to say that for a clear increase of KE one need a drastic increase of α.

[20]Speaking well English rises the human capital on the international labor market significantly with all its sequences, resulting in higher lifetime income.

Altuna, O. and Basurto, A. (2013) *A Guide to Language Use Observation. Survey Methods.* Soziolinguistika Klusterra.

Blume, A. and Board, O. (2013) "Language Barriers", *Econometrica*, Vol. 81, No. 2 (March, 2013), 781-812.

Chen, M.K. (2013) "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets" *American Economic Review* 2013, 103(2): 690-731

Church, J. and King, I. (1993) "Bilingualism and Network Externalities", *Canadian Journal of Economics* 26, no. 2, pp. 337-345.

Crystal, D. (1987) *The Cambridge Encyclopedia of Language*, Cambridge, UK: Cambridge University Press.

Fishman, J.A. (2001) "Why is it so Hard to Save a Threatened Language?" in J.A. Fishman (Ed.) *Can Threatened Languages be Saved*? Clevedon: Multilingual Matters.

Gabszewicz, J., Ginsburgh, V. and Weber, S. (2011a) "Bilingualism and Communicative Benefits", *Annales d'Economie et Statistique.* no. 101-2, pp. 271-86.

Ginsburg, V., Ortuño-Urtín, I. and Weber, S. (2007) "Learning Foreign Languages: Theoretical and Emprical Implications of the Selten and Pool Model", Journal of Economic Behavior and Organization, vol. 64, pp. 337-347.

Ginsburgh, V. and Weber, S. (2011) *How Many Languages Do We Need? The Economics of Linguistic Diversity.* Princeton, NJ: Princeton University Press.

Grin, F., Sfreddo, C. and Vaillancourt, F. (2010) *The Economics of the Multilingual Workplace.* New York, NY: Routledge.

Härdle, W. and Mammen, E. (1993) "Comparing Nonparametric Versus Parametric Regression Fits", *Annals of Statistics* 21: 1926-1947.

Jones, H.M. (2012) *A statistical overview of the Welsh language.* Welsh Language Board, Cardiff.

Iriberri, N. and Uriarte, J.R. (2012) "Minority Language and the Stability of Bilingual Equilibria", *Rationality and Society* 2: 442-462.

Mira, J.B. and Paredes, A. (2005) "Interlinguistic Similarity and Language Death Dynamics", *Europhysics Letters* 69: 1031-1034.

Kahneman, D. and Tversky, A.(1979) "Prospect Theory: An Analysis of Decision under Risk", *Econometrica* 47: 263-289.

Patriarca, M. and X. Castelló, Uriarte, J.R., Eguiluz, V.M. and San Miguel, S. (2012) "Modeling Two Language Competition Dynamics" *Advances in Complex Systems,* Vol. 15, Nos. 3 & 4:1250048-1-24

Rubinstein, A. (2000) *Economics and Language*, Cambridge:CUP

Shapiro, D. and Stelcner, M. (1997) "Language and Earnings in Quebec: Trends over-Twenty Years, 1970-1990." Canadian Public Policy, vol. 2, no. 23, pp. 115-40.

Selten R. and Pool, J. (1991) "The Distribution of Foreign Language Skills as a Game Equilibrium" in R. Selton ed: *Game Equilibrium Models IV: Social and Political Interaction*, Berlin: Springer-Verlag, pp. 64-87.

UNESCO (2002). "UNESCO Universal Declaration On Cultural Diversity"

Vaillancourt, F. (1980) "Differences in Earnings by Language Group in Quebec 1970: An Economic Analysis". Quebec: International Center for Research on Bilingualism.

Weibull, J.W. (1995) *Evolutionary Game Theory,* Cambridge, Mass.: the MIT Press

7 Appendix

7.1 Nonparametric Estimation procedure

Given a sample $\{\alpha_i, KE_i\}_{i=1}^{n}$ one wants to estimate the conditional expectation $E[KE|\alpha] = m(\alpha)$ under the assumption that $m(\cdot)$ is a smooth function having third order Lipschitz continuous derivatives. The errors $v = KE - m(\alpha)$ have finite variance. One may add some conditions on the distribution of α if one wants to calculate the statistical properties of the now described estimator: For a weight or kernel function $K(\cdot)$ for which we chose the Epanechnikov kernel $K(u) = 0.75 \cdot (1-u^2)_+$ (the subindex $+$ indicates that the function is set to zero if $1-u^2$ is negative) and bandwidth h_x we take

$$\widehat{m(x)} = \underset{m,m_1,m_2}{argmin} \sum_{j=1}^{n} \left(KE_j - m - m_1 \cdot (\alpha_j - x) - m_2 \cdot (\alpha_j - x)^2\right)^2 K(\frac{\alpha_j - x}{h_x}) \tag{7}$$

as an estimate for $m(x)$. This is the well-known local quadratic kernel estimator. Letting x run over the range of α (here simply over all sample observations α_i) we can draw than the function estimate of $m(\cdot)$ which is compared than with our model for PKE.

7.2 The bootstrap test of Härdle and Mammen (1993)

We want to check the null hypothesis that the parametric model does not significantly deviate from the nonparametric fit which is supposed to reflect the true model but with a potential smoothing bias. The proposed test statistic is

$$T_{ct} = \frac{1}{n_{ct}} \sum_{i=1}^{n_{ct}} \left(\widetilde{\widehat{PKE}}_{cti} - \widehat{m(\alpha_{cti})}\right)^2 , \tag{8}$$

where $\widehat{m(\alpha_{cti})}$ is the nonparametric data fit of KE on α. Let $\widehat{PKE}_{cti}$ be the parametric prediction along our theoretical model. To avoid potential smoothing bias problems, it is recommended to let it pass through the kernel smoother, too. That is, estimation procedure (7) is applied to $\{\alpha_{cti}, \widehat{PKE}_{cti}\}_{i=1}^{n_{ct}}$, and call the results $\widetilde{PKE}_{cti}$. To simulate the p-value for test statistic T_{ct} under the null hypothesis one applies wild bootstrap. That is, we keep the α_{cti} but generate new responses by $\widetilde{KE}_{cti} = \widehat{PKE}_{cti} + (KE_{cti} - \widehat{PKE}_{cti}) \cdot N(0,1)$ (i.e. take the parametrically prediction and add a new normal random term respecting potential heteroscedasticity). Then we calculate the test statistic from this new sample which in fact has been generated under the null hypothesis. This can be done for example a 100 times. The percentage of these statistics being larger than the original one (8) is a simulated approximate of the p-value of our test.