DISCUSSION PAPER SERIES

No. 7587

MARKETS AND LINGUISTIC DIVERSITY

Ramon Caminal

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP7587.asp**

ISSN 0265-8003

MARKETS AND LINGUISTIC DIVERSITY

Ramon Caminal, Institut d'Anàlisi Econòmica CSIC and CEPR

Discussion Paper No. 7587
December 2009

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **INDUSTRIAL ORGANIZATION**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Ramon Caminal

CEPR Discussion Paper No. 7587

December 2009

ABSTRACT

Markets and linguistic diversity

Producers of cultural goods and media products can only make their specific contents available to their audiences and readerships through a particular language. The choice of language is a trivial decision if consumers are monolingual. However, the fraction of bilingual consumers is high in some areas and rising everywhere because of the rapid expansion of English as a second language. In this paper I argue that, independently of the gains associated with the use of a lingua franca, the very existence of bilingual consumers may seriously bias market outcomes against minority languages. In particular, I show that the level of linguistic diversity determined by profit maximizing firms tends to be inefficiently low, except when and where the cost of producing a second linguistic version becomes sufficiently low. Thus, the model provides an efficiency argument supporting policies that protect minority languages in these markets.

JEL Classification: D43, L13 and L82
Keywords: language, product variety and translation

Ramon Caminal
Institut d'Anàlisi Econòmica, CSIC
Universitat Autonoma de Barcelona
Bellaterra
E-08193 Barcelona
SPAIN

Email: caminal@cc.uab.es

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=109343

Submitted 27 November 2009

Markets and linguistic diversity

October 2009

Abstract

Producers of cultural goods and media products can only make their specific contents available to their audiences and readerships through a particular language. The choice of language is a trivial decision if consumers are monolingual. However, the fraction of bilingual or multilingual consumers is high in some areas and rising everywhere because of the rapid expansion of English as a second language. In this paper I argue that, independently of the gains associated with the use of a *lingua franca*, the very existence of bilingual consumers may seriously bias market outcomes against minority languages. In particular, I show that the level of linguistic diversity determined by profit maximizing firms tends to be inefficiently low, except when and where the cost of producing a second linguistic version becomes sufficiently low. Thus, the model provides an efficiency argument supporting policies that protect the presence of minority languages in these markets.

JEL Classification numbers: D43, L13, L82

Key words: Product variety, language, translation

1 Introduction

Producers of cultural goods (like books and films) and media products (newspapers, TV and radio) can only make their specific contents available to their audiences and readerships through a particular language. The choice of language is a trivial decision if consumers are monolingual. However, in many local markets, as well as in the global market, a substantial fraction of consumers are competent in more than one language. It is precisely the presence of multilingual consumers that makes language selection a difficult decision

for firms and a sensitive issue for the general public. Let us first discuss this topic in the context of the global market. It is well known that the expansion of English as a second language has accelerated over the last decades.[1] Some commentators have expressed their concern about the potential negative effect of this expansion on the presence and diffusion of other languages. For now the signals are weak but non-negligible. For instance, US and UK-based TV stations specialized in international news are attracting larger audiences around the world. At the same time stations based in non-English speaking countries have set up English channels (Al Jazeera, Russia Today, France 24)[2]. Films and books originally made in English already enjoy a clearly dominant position in world markets, although they are typically translated or dubbed into local languages. However, incentives to pay the costs of translation and dubbing may be significantly reduced as consumers' competence in English is enhanced.

Clearly, both the expansion of English as a second language and the integration of cultural goods and media markets is likely to speed up in the coming years and will probably convey very substantial benefits. However, they may also involve significant costs and market failures that we need to pay attention to. In particular, the reduction in the degree of linguistic diversity may specifically harm monolingual social groups; but more generally, it may negatively affect those consumers whose mother tongue is not English and have a preference for consuming these products in a different language. Moreover, the presence of a particular language in the media and cultural goods markets is crucial for its vitality and prestige, potentially influencing its medium and long-term dynamics.

Geographic areas where the majority of the population is bilingual can provide useful insights on the long-run implications of the expansion of English. Take the example of Catalonia. A very large fraction of the 7 million inhabitants can speak and read the two main languages: Catalan and Spanish (Castilian). Surveys conducted over the last twenty years indicate that

[1] According to The Economist (Dec 13th 2006) nearly a quarter of the world's population speaks some English. That includes those who speak it as their mother tongue (400 million), and those who speak it fluently as their second language (another 400 million). It is also estimated that about a billion are learning it. Graddol (2006) argues that the last figure is likely to double within a decade.

[2] The incentives to adopt English are probably stronger on the internet. A prominent example is the electronic version of Der Spiegel, Spiegel Online International. Some major European newspapers are also considering similar moves.

Spanish is the family language for roughly half of the population, while Catalan is for the other half (data on daily use of the two languages also show approximately the same fifty-fifty pattern.)[3] One could naively expect that one half of consumption of cultural goods and media products in Catalonia would be in Catalan and the other half in Spanish. But this is not the case, especially if we focus on non-subsidized, privately provided goods and services. More specifically, 24% of the TV audience consumes programs in Catalan, but only a tiny fraction is broadcast by private stations. Similarly, 43% of the radio audiences correspond to programs in Catalan, but the fraction that is supplied by private radio stations is also small. Private supply in Catalan is higher in the newspaper and book markets. About 22% of the newspapers and 20% of books (excluding textbooks) consumed in Catalonia are in Catalan. Finally, the consumption of films dubbed into Catalan or originally produced in this language is close to zero (subtitles are rarely used in Spain).[4] It is important to emphasize that only a few extra million consumers outside Catalonia speak or read Catalan, while in most of these examples the relevant market is either Spain (46 million, all competent in Spanish) or the world's Spanish speaking population (approximately, 400 millions.)

These indicators can be interpreted in different ways. Nevertheless, the low private provision of products in Catalan, combined with the wide political support for the public financing of TV and radio stations that broadcast programs exclusively in Catalan[5], suggest that market outcomes might be biased against minority languages.

As suggested above, the relationship between linguistic preferences and market outcomes is dynamic with causality running both ways. This paper contributes to our understanding of these issues by focusing on the causality from preferences to outcomes. In particular, I ask whether or not markets

[3] See, for instance, Generalitat de Catalunya (2003), Les estadístiques d'usos lingüístics. The most recent surveys (Les estadístiques d'usos lingüístics, 2008) indicate that the fraction of the Catalan-oriented population has fallen below 50%, probably as a result of recent immigration flows. However, the fifty-fifty distribution is still a good approximation for consumers of books, newspapers and films (Baròmetre de la Comunicació i la Cultura, 2008).

[4] These indicators have been provided by industry associations and audience monitoring agencies. The figures given in the text are averages over the period 1999-2008.

[5] The charter of the Catalan TV and radio networks was approved in the regional parliament by unanimity in 1983. It was reformed in 2007 with the support of 87% of the members of the parliament.

tend to provide too little linguistic diversity, for a given distribution of linguistic preferences. Thus, I analyze a static model of product variety where consumers have preferences over the intrinsic characteristics (content) as well as the language of the product.[6] The model is agnostic about the origin of linguistic tastes. A preference for a particular language may simply reflect higher proficiency or, on the contrary, the desire to learn the language. It could also emerge from ideological, or even esthetic, reasons. Consumers belong to distinct linguistic communities of unequal size. These communities are not completely segmented, in the sense that members of the minority are bilingual, while the majority members may be bilingual (symmetric bilingualism) or monolingual (asymmetric bilingualism). Bilingual consumers are not indifferent about the language of the product; they strictly prefer to consume products in one of the languages (say, their mother tongue), although they may be willing to consume products in their second language depending on availability, content preferences, and price differentials.

The model extends the spokes framework (Chen and Riordan, 2007) by adding an additional dimension of product differentiation (language). Thus, consumers may trade off a good match in terms of content against a good linguistic match. In fact, a key parameter of the model is the degree of linguistic substitutability relative to the substitutability of contents. It is very convenient to start off by assuming that each variety (defined by its content) must be supplied in a single language (the adoption model). Next, the analysis is extended by allowing firms to provide more than one linguistic version of each variety (the translation model). In the real world, translations are not restricted to books and films, but nowadays they are somewhat present in most segments of cultural goods and media markets.

The main results of the paper are the following. Firstly, in the adoption model market forces are biased against minority languages; that is, the fraction of varieties in the minority language supplied in equilibrium is below the level that maximizes total surplus (the level of linguistic diversity is inefficiently low). Let us consider the social planner's problem. If the degree of linguistic substitutability is not too high (with respect to the size of the minority community) then it is efficient to supply a positive fraction of varieties in the minority language, since a fraction of consumers can be

[6] There is a huge literature on optimal product diversity. The most significant milestones include Dixit and Stiglitz (1977), Salop (1979), Mankiw and Whinston (1986) and Chen and Riordan (2007). These are models where product differentiation is unidimensional.

assigned to products on the basis of their linguistic preferences, which more than compensates the content mismatch. In fact, as the degree of linguistic substitutability falls (i.e., the intensity of linguistic preferences increases), or as the size of the minority community increases, then the optimal fraction of varieties in the minority language increases. However, private incentives to supply goods in the minority language are smaller than social incentives. The main driving force of this result is the size of the consumer base. Actually, this is the only force in a regulated environment where prices are fixed exogenously (and independently of the language of the product). In this case, no firm would ever find it profitable to supply its product in the minority language, since that would imply lower sales. In an unregulated environment, however, some firms may be willing to specialize, adopt the minority language, and take advantage of the higher willingness to pay of members of the minority linguistic community. However, a firm that switches to the minority language is unable to capture all the surplus generated.

The second main result concerns the translation model. In this case, the market outcome may exhibit insufficient or excessive linguistic diversity. However, unlike most models of product differentiation, the sign of the inefficiency can be associated to the value of observable variables. In particular, excessive diversity arises only in markets where the cost of translation is so low that the fraction of products in the minority language is disproportionately high with respect to the relative size of the linguistic minority. In order to understand this result it is important to note that in the extreme symmetric case where both linguistic communities have the same size, despite of the fact that the level of linguistic diversity is always efficient, the market allocation is inefficient because firms tend to generate too many translations. The reason is that most consumers attracted by a second linguistic version of a particular variety come from rival firms (business stealing effect). Thus, back in the asymmetric case, the ambiguous result is a combination of the forces described in the adoption model and the excessive incentives to translate. In particular, if the cost of translation is relatively high then we are very close to the adoption model. In contrast, if the cost of translation is sufficiently low then in equilibrium most firms offer a second linguistic version in order to steal consumers from rival firms.

Summarizing, the adoption model formalizes the widespread perception that market forces tend to work against minority languages. However, the efficiency of market outcomes is likely to improve as the application of new technologies reduces the costs of translations. If the cost reduction is suffi-

ciently drastic then the sign of the inefficiency can even be reversed.

There is a growing literature on the economics of language, which has examined a broad range of issues, such as the acquisition of a second language (Selten and Pool, 1991; Church and King, 1993, and Lazear, 1999), the intergenerational transmission of mother tongues (John and Yi, 2001), the role of language in foreign trade (Melitz, 2007b), the choice of official languages in multilingual societies (Ginsburgh et al., 2005), and the relationship between language policy and human capital accumulation (Ortega and Tangeras, 2008). To the best of my knowledge the issue of language adoption in cultural goods and media markets has not yet been the subject of formal economic analysis.

More closely related to the present paper is the work about the impact of English dominance on literary translation (Melitz, 2007a; Ginsburgh et al., 2007.) These papers consider a very different framework with completely segmented markets and monolingual consumers.

The next section presents the adoption model. Section 3 is devoted to the translation model. Some concluding remarks close the paper.

2 Language adoption

2.1 A model of symmetric bilingualism

The model builds on the spokes framework recently presented by Chen and Riordan (2007), which provides a (spatial) representation of consumer preferences over an arbitrary number of differentiated products. I interpret such a preference space as referring to the intrinsic characteristics (content) of cultural goods and media products. On the top of this, I introduce an additional dimension of product differentiation: the language of the product. I restrict myself to the case that the number of active firms is equal to the number of potential varieties.[7] Also, in order to enhance tractability, I use the continuous approximation of the spokes framework proposed by Caminal and Granero (2008). Thus, the fraction of products supplied in a particular language is a continuous variable.

[7]If the number of firms is endogenously determined by a zero profit condition then in equilibrium there may be excessive or insufficient entry. Here, we are exclusively concerned with linguistic diversity and therefore it is very convenient to keep the diversity of intrinsic characteristics fixed. In Section 4 I discuss the implications of this assumption.

More specifically, the model considers a continuum of firms (mass one) and consumers (also, mass one.) Each firm produces a differentiated product, which can be supplied in one of two possible languages, S and C. In the next section I allow firms to supply two linguistic versions of the good, one in S and one in C.

Each consumer has a preference for only two varieties (defined in terms of intrinsic characteristics) and consumes one unit of one of these two varieties. In fact, each consumer's identity is given by three elements: the pair of selected varieties, the relative preference for these two varieties (location in the $[0,1]$ segment) and the preferred language. Let us examine these three elements sequentially.

Consumers are uniformly distributed over all possible pairs of varieties. This implies, in particular, that the subset of consumers that have a preference for a particular variety are also uniformly distributed over their second variety. Thus, each firm has a negligible effect on the demand faced by any other individual firm (monopolistic competition.)

A fraction α of all consumers prefer to buy goods supplied in the C-language and $1-\alpha$ in the S-language. I will refer to the first group as C-consumers and to the second group as S-consumers. I assume that $\alpha \in \left(0, \frac{1}{2}\right)$, and hence we call C and S the minority and majority language, respectively. Another important modeling choice concerns the correlation between language and non-language (content) preferences. In the limiting case of perfect correlation, C-consumers would be exclusively interested in a fraction α of all possible varieties and S-consumers in the remaining $1-\alpha$. In this case the language adoption decision would be trivial. Alternatively, I will focus on the opposite extreme scenario and assume zero correlation between preferences about language and non-language characteristics (and zero correlation with respect to relative preferences.) Thus, all firms face a customer base with the same distribution of linguistic preferences.

Consumers with a preference for varieties i and j are uniformly distributed in the interval $[0,1]$, which represents the intensity of their relative preferences for their intrinsic characteristics. Thus, a consumer located at $x \in [0,1]$, obtains a utility of $R - x - \lambda\Theta$ if she consumes one unit variety i and $R-(1-x)-\lambda\Theta$ if she consumes one unit of variety j. In the jargon of standard spatial models, I assume linear transportation costs and normalize the unit cost to one. Variable Θ takes value 0 if the product is supplied in the most preferred language and 1, otherwise. Thus, $\lambda > 0$ represents the reduction in utility associated to the use of the least preferred language

(language mismatch).

I assume that $R - \lambda > 2$, in order to guarantee that in equilibrium all consumers are served, and that firms are subject to constant and identical marginal costs, which for simplicity are normalized to zero. Finally, I restrict the analysis to the case $\lambda < 2$, otherwise firms supplying goods in different languages would not compete with each other.

2.2 The first best

Let us denote by β the fraction of varieties supplied in C. For a given β, there is a fraction β^2 of consumers whose two selected varieties are supplied in C, a fraction $2\beta(1-\beta)$ with one variety in S and one in C, and a fraction $(1-\beta)^2$ with both varieties in S.

For those consumers facing two varieties with the same language then efficiency requires the allocation of consumers to the closest variety. Thus, the total surplus that can be obtained by consumers with access to two C-varieties and two S-varieties is $R - \frac{1}{4} - (1-\alpha)\lambda$ and $R - \frac{1}{4} - \alpha\lambda$, respectively. The average transportation costs is in both cases $\frac{1}{4}$, and the only difference is the weight of the language mismatch.

The optimal allocation of consumers with access to two varieties supplied in different languages will obviously depend on their language preferences. Without loss of generality suppose that the variety supplied in C is located at zero and the variety supplied in S is located at one. Thus, C-consumers should consume the variety located at zero, if and only if $x \leq \min\left\{\frac{1+\lambda}{2}, 1\right\}$. If $\lambda < 1$ then $\frac{1+\lambda}{2} < 1$ and those C-consumers close to 1 must consume the variety supplied in S (the linguistic mismatch is dominated by the intensity of preferences over intrinsic characteristics.) However, if $\lambda \geq 1$ it is optimal to allocate consumers exclusively according to their linguistic preferences. Thus, the maximum surplus obtained in those segments with one variety in each language, $A(\lambda)$, is given by $R - \frac{1}{4} - \frac{\lambda}{2} + \frac{\lambda^2}{4}$ if $\lambda < 1$, and $R - \frac{1}{2}$ otherwise.

Thus, total welfare, W, can be written as:

$$W(\beta) = \beta^2\left[R - \frac{1}{4} - (1-\alpha)\lambda\right] + 2\beta(1-\beta)A(\lambda) + (1-\beta)^2\left[R - \frac{1}{4} - \alpha\lambda\right]$$

Thus, if $\lambda < 1$ the optimal fraction of varieties in the minority language, β^*, is:

$$\beta^* = \max\left\{0, \frac{1}{2} - \frac{1-2\alpha}{\lambda}\right\}$$

First, note that $\beta^* > 0$ if and only if $\lambda \geq 2(1-2\alpha)$. Second, $\beta^* < \alpha$. Third, β^* increases with both α and λ. In order to gain some intuition let us consider the case $\beta = 0$. If variety i switches from S to C, then all consumers who have a taste for variety i, face a choice between a variety in C and a variety in S. In this case, because of the linguistic switch S-consumers in average will loose $\lambda\left(1 - \frac{\lambda}{2}\right) < \lambda$, since some of them will switch to their alternative variety (supplied in S). In contrast, C-consumers will gain in average $\lambda\left(1 + \frac{\lambda}{2}\right) > \lambda$, since more consumers will be induced to purchase variety i. Thus, only if α is sufficiently high, $\alpha\,\lambda\left(1 + \frac{\lambda}{2}\right) > (1-\alpha)\,\lambda\left(1 - \frac{\lambda}{2}\right)$, it is efficient to supply variety i in S. Also, as α and λ increase the surplus associated to supplying varieties in C also increases.

If $\lambda \geq 1$,

$$\beta^* = \max\left\{0, \frac{\alpha\lambda - \frac{1}{4}}{\lambda - \frac{1}{2}}\right\}$$

Thus, $\beta^* > 0$ if and only if $\lambda > \frac{1}{4\alpha}$. The area of parameter values for which $\beta^* > 0$ is depicted in Figure 1.

2.3 Market equilibrium

In the market game firms will adopt a particular language depending on its relative profitability. In a regulated environment where prices cannot vary with the language, adopting the minority language implies lower sales and hence it cannot be profitable. Instead, if prices are endogenous, an individual firm may find it optimal to adopt the minority language, charge a higher price, and exploit the higher willingness to pay of C-consumers.

The specific details of the equilibrium do depend on firms' ability to price discriminate between members of different language communities. If the linguistic composition of consumers exhibits sufficient regional variation, then firms have incentives to set different prices in different regions. For convenience, I present first the perfect discrimination case (geographically segmented linguistic communities.) In Section 2.4 I discuss the case of no price discrimination.

Given that each individual firm has a negligible influence on other firms' decisions, it does not matter whether language adoption and price decisions are taken sequentially or simultaneously.

Let $p^k(l)$ be the price charged by a variety supplied in the $k-$language to the $l-$community, $k, l = C, S$. Consider, for instance, the price charged to S-consumers by firm i supplying the good in S. A fraction β of these consumers have a variety in C as an alternative choice, while a fraction $(1-\beta)$ have a variety in S. If we denote firm i's price by p_i then the profit function can be written as:

$$\pi_i = p_i \left\{ \beta \left[1 + \lambda + p^C(S) - p_i \right] + (1-\beta) \left[1 + p^S(S) - p_i \right] \right\}$$

provided that $1 + \lambda + p^C(S) - p_i \leq 2$.

If we evaluate the first order condition of an interior solution at $p_i = p^S(S)$, then the "joint" reaction function is given by:

$$p^S(S) = \frac{1 + \beta \left[\lambda + p^C(S) \right]}{1 + \beta}$$

The other three optimization problems provide three more equations. By solving the system we obtain the candidates to (symmetric, pure strategy) equilibrium prices:

$$p^S(S) = 1 + \frac{\lambda\beta}{2}$$

$$p^C(S) = 1 - \frac{\lambda(1-\beta)}{2}$$

$$p^C(C) = 1 + \frac{\lambda(1-\beta)}{2}$$

$$p^S(C) = 1 - \frac{\lambda\beta}{2}$$

We still need to check that, given these prices, no firm wishes to deviate. It turns out that if λ is sufficiently close to 2 (the threshold is a decreasing function of β) then a firm supplying the product in S finds it optimal to charge C-consumers a price higher than $\left(1 - \frac{\lambda\beta}{2}\right)$ and serve only those consumers who only have access to two varieties in S. Similarly, a firm supplying the product in C finds it optimal to charge S-consumers a price higher

than $\left(1-\frac{\lambda(1-\beta)}{2}\right)$ and serve only those consumers who only have access to two varieties in C. It turns out that for these parameter values there is no symmetric equilibrium in pure strategies.

Note that a firm that supplies its product in a particular language charges a higher price to those consumers that prefer that language: $p^k(k) > p^k(l), k = C, S$, and $l \neq k$. In fact, the price differential, $p^k(k) - p^k(l)$, decreases with the fraction of products supplied in k.

The firm's profits depend on the language adopted and the language distribution:

$$\pi^S(\beta) = \alpha\left(1-\frac{\lambda\beta}{2}\right)^2 + (1-\alpha)\left(1+\frac{\lambda\beta}{2}\right)^2$$

$$\pi^C(\beta) = \alpha\left(1+\frac{\lambda(1-\beta)}{2}\right)^2 + (1-\alpha)\left(1-\frac{\lambda(1-\beta)}{2}\right)^2$$

In equilibrium $\pi^S(\beta^e) = \pi^C(\beta^e)$, provided $\beta^e > 0$, which implies that:

$$\beta^e = \max\left\{0, \frac{1}{2} - \frac{2(1-2\alpha)}{\lambda}\right\}$$

Note that $\beta^e > 0$ if and only if $\lambda > 4(1-2\alpha)$. Figure 1 shows that the area of parameter values for which $\beta^e > 0$ is smaller than than the area corresponding to $\beta^* > 0$. More generally, we can compare the equilibrium allocation with the first best (Figure 2 draws β^e and β^* for a given value of λ):

Proposition 1 *The equilibrium of the adoption model exhibits insufficient linguistic diversity, in the sense that the fraction of varieties in the minority language is inefficiently low; i.e., $\beta^e \leq \beta^*$, and provided $\beta^* > 0$ then $\beta^e < \beta^*$.*

In order to gain some intuition about the discrepancy between social and private incentives, let us consider the case $\beta = 0$. Equilibrium prices are $p^S(S) = p^S(C) = 1, p^C(S) = 1 - \frac{\lambda}{2}, p^C(C) = 1 + \frac{\lambda}{2}$. If a particular firm considers switching from S to C, then this would cause an upward shift in the demand by C-consumers and a downward shift in the demand by S-consumers. Figure 3 reflects the shift in demand by C-consumers, drawn for the case $\lambda < 1$. The shaded area, $A + B + C + D$, represents the increase in total welfare. However, the firm cannot appropriate all this extra surplus

because it faces two conflicting goals and has only one instrument: the price. On the one hand, the firm adopting C would like to extract the extra surplus from their previous customers. On the other hand, it would like to attract new consumers (business stealing). If the firm charges $p^C(C) = 1 + \lambda$, then total sales remain unchanged, and the firm captures all the extra surplus from existing customers but it attracts no new consumer. In this case, the extra profits are equivalent to area A. If instead the firm charges $p^C(C) = 1$, then total demand has increased by an amount equal to λ. In this case, it attracts an efficient amount of new consumers but it does not obtain any extra profits from existing customers. In this case, the total amount of extra profits would be given by areas $E+F$, which is equivalent to A. The optimal price is an intermediate one, $p^C(C) = 1 + \frac{\lambda}{2}$, which reflects the optimal balance between these two conflicting goals. Maximum profits are given by areas $E + F + C = A + C$. Thus, the firm is unable to capture the entire surplus: it misses $B + D$.

The argument for S-consumers is analogous: the profit loss is larger than the loss in total welfare.

2.4 Discussion

2.4.1 No price discrimination

The assumption that firms can perfectly discriminate between members of different linguistic communities was very convenient, but it may not be a good approximation in some real world examples. In the Appendix I analyze the opposite extreme case where firms charge the same price to all consumers. It turns out that firms adopting C are harmed more intensively if they cannot price discriminate, and as a result incentives to adopt C are further reduced, which exacerbates the market bias against minority languages.[8]

2.4.2 Asymmetric bilingualism

In the benchmark model I assumed that members of both linguistic communities are bilingual in the sense that they are competent in both languages, although every consumer has a preference for consuming the product in one of the languages. In some real world examples a fraction of consumers are

[8]The set of parameter values for which there is no symmetric equilibrium in pure strategies is also higher than in the case of price discrimination.

monolingual and as a result language choices affect aggregate sales. An important observation is that members of small language communities tend to be competent in at least a second language, while the proportion of monolingual consumers is higher in large language communities. Thus, it makes sense to consider an alternative specification of the present model where C-consumers are bilingual and have the preferences described in the benchmark model (they experience a utility loss λ if they consume the good in S) but S-consumers are monolingual and hence experience an infinite utility loss if they consume the good in C.

In the Appendix I analyze both the first best allocation and the market equilibrium of this version of the model. I show that both β^* and β^e are lower than in the benchmark model, which is a very intuitive result. What is more important, is that the main insights brought about by Proposition 1 remain unchanged: $\beta^* \geq \beta^e$, and if $\beta^* > 0$ then $\beta^* > \beta^e$ (market outcomes are biased against minority languages).

2.4.3 Advertising

Suppliers of some media products (TV, radio) charge a zero price to consumers. Instead, they obtain revenue from advertising. Does the financing channel affect the language choice? Let us consider a version of the present model where firms charge a zero price but they obtain an advertising revenue A. Consumers' net utility is a decreasing function of the intensity of advertising, $\eta(A)$, where $\eta' < 0, \eta'' < 0$.[9] Thus, instead of paying a price consumers experience a disutility associated with advertising. In order to attract consumers, firms can reduce the intensity of advertising instead of cutting the price (in this case it is not feasible to discriminate between members of different linguistic communities). It turns out the main qualitative properties of Proposition 1 remain unchanged with respect to the financing channel (See Appendix.) A firm that adopts the minority language will tend to increase the intensity of advertising trying to exploit C-consumers' higher willingness to pay, but again the ability to appropriate a sufficient amount of surplus using this channel is very limited.

[9] See, for instance, Peitz and Valletti (2008) for an analogous comparison between pay-tv versus free-to-air.

2.4.4 Second best

The main result of this section is that market forces tend to deliver insufficient linguistic diversity. This result provides a possible justification for government intervention based purely on efficiency grounds. A possible instrument would be production subsidies to suppliers of goods in minority languages. However, in order to implement the first best, we would need to design those subsidies in a way that they eliminate the price distortion associated to language choices. A much simpler policy would consist of subsidies conditional only on language choices. In order to justify such policy we would need to compare equilibrium outcomes with a second best scenario, in which total welfare is computed conditional on the distortionary price behavior predicted in equilibrium. If we let β^{**} be the fraction of varieties supplied in the minority language, then (see Appendix) $\beta^{**} \geq \beta^{e}$, and the inequality is strict for those parameter values such that $\beta^{**} \geq 0$. In other words, subsidies to firms producing the good in the minority language can improve the efficiency of market outcomes.

3 Translations

Suppose now that each firm can choose one of the following three options: (i) supply the product in S, (ii) supply the product in C, and (iii) supply two linguistic versions of the product, one in S and one in C, and pay an extra cost $F > 0$.[10]

Let us now denote by β_s, β_c the fractions of goods supplied exclusively in S and C, respectively. Therefore, $1-\beta_s-\beta_c$ is the fraction of goods supplied in both languages (translations.)

3.1 The first best

The fraction of varieties available in S and C is $1-\beta_c$ and $1-\beta_s$, respectively. As in the previous section, the optimal allocation of consumers in those segments where the two varieties are supplied in a different language will depend on whether λ is higher or lower than 1. In the Appendix I present

[10] Books and films are typically created in a particular language and then translated or dubbed into other languages. It might be argued that something is lost in the process of translation. Here, we disregard this possibility by treating both languages in a symmetric fashion.

the case $\lambda > 1$, but in the text I restrict attention to the case $\lambda \leq 1$. A fraction $(1-\beta_c)^2$ of S-consumers have access to two varieties in S and hence the average surplus is $R - \frac{1}{4}$. A fraction β_c^2 have access to two varieties in C and the average surplus is $R - \lambda - \frac{1}{4}$. Finally, a fraction $2\beta_c(1-\beta_c)$ have access to one variety in each language and the average surplus taking into account the optimal allocation of those consumers is $R - \frac{1}{4} - \frac{\lambda}{2} + \frac{\lambda^2}{4}$. The expressions are analogous for $C-$consumers. Thus, total welfare is:

$$W(\beta_c, \beta_s) = R - \frac{1}{4} - \alpha\lambda\left[\beta_s(1-\beta_s)\left(1-\frac{\lambda}{2}\right) + \beta_s^2\right] -$$

$$-(1-\alpha)\lambda\left[\beta_c(1-\beta_c)\left(1-\frac{\lambda}{2}\right) + \beta_c^2\right] - F(1-\beta_s-\beta_c)$$

where $\beta_s + \beta_c \leq 1$. If we denote by $z \equiv \frac{F}{\alpha\lambda}$ then the candidates to be the optimal values of β_s and β_c are given by the first order conditions of an interior solution:

$$\widetilde{\beta}_s(z) = \begin{cases} 0, \text{ if } z \leq 1 - \frac{\lambda}{2} \\ \frac{1}{2} + \frac{z-1}{\lambda}, \text{ if } 1 - \frac{\lambda}{2} \leq z \leq 1 + \frac{\lambda}{2} \\ 1, \text{ if } z \geq 1 + \frac{\lambda}{2} \end{cases}$$

$$\widetilde{\beta}_c(z) = \begin{cases} 0, \text{ if } \frac{\alpha z}{1-\alpha} \leq 1 - \frac{\lambda}{2} \\ \frac{1}{2} + \frac{\frac{\alpha z}{1-\alpha}-1}{\lambda}, \text{ if } 1 - \frac{\lambda}{2} \leq \frac{\alpha z}{1-\alpha} \leq 1 + \frac{\lambda}{2} \\ 1, \text{ if } \frac{\alpha z}{1-\alpha} \geq 1 + \frac{\lambda}{2} \end{cases}$$

It turns out that $\widetilde{\beta}_s(z) + \widetilde{\beta}_c(z)$ is sometimes higher than 1 (the constraint is binding). It will be useful to consider Condition A :

$$\frac{1-\alpha}{\alpha}\left(1-\frac{\lambda}{2}\right) > 1 + \frac{\lambda}{2}$$

i.e.,

$$\alpha < \frac{1}{2}\left(1-\frac{\lambda}{2}\right)$$

Note that for those parameter values that satisfy Condition A the adoption model prescribes $\beta^* = 0$.

If Condition A holds then for all values of z such that $\widetilde{\beta}_c(z) > 0$ we have that $\widetilde{\beta}_s(z) = 1$. Therefore, in this case the optimal values are $\beta_c^* = 0$ and

$\beta_s^* = \widetilde{\beta}_s(z)$. That is, if $z \geq 1 + \frac{\lambda}{2}$, then all varieties are supplied exclusively in S. However, as z falls below $1 + \frac{\lambda}{2}$ an increasing number of varieties is supplied in both languages.

If Condition A fails then it is possible to have both β_s^* and β_c^* strictly positive. There are two possible regions. If $z \geq (2 - 2\alpha)$ then $\beta_s^* = \widetilde{\beta}_s(2 - 2\alpha)$ and $\beta_c^* = \widetilde{\beta}_c(2 - 2\alpha)$. That is, if translation costs are sufficiently high then there are no translations and language choices coincide with those of the adoption model. Alternatively, if $z < (2 - 2\alpha)$, then $\beta_s^* = \widetilde{\beta}_s(z)$ and $\beta_c^* = \widetilde{\beta}_c(z)$.

It is immediate to check that: (i) $\beta_s^* \geq \beta_c^*$, (ii) neither β_s^* nor β_c^* decrease with F, and (iii) both β_s^* and β_c^* may increase or decrease with λ. The economic intuitions are straightforward.

3.2 Market equilibrium

The pricing game is analogous to the one discussed in Section 2.[11] Thus, equilibrium prices are given by:

$$p^S(S) = 1 + \frac{\lambda \beta_c}{2}$$

$$p^C(S) = 1 - \frac{\lambda(1 - \beta_s)}{2}$$

$$p^C(C) = 1 + \frac{\lambda \beta_s}{2}$$

$$p^S(C) = 1 - \frac{\lambda(1 - \beta_c)}{2}$$

Profits from supplying the product in S and C exclusively are given respectively by:

$$\pi^S(\beta_s, \beta_c) = \alpha \left[p^S(C)\right]^2 + (1 - \alpha)\left[p^S(S)\right]^2$$

$$\pi^C(\beta_s, \beta_c) = \alpha \left[p^C(C)\right]^2 + (1 - \alpha)\left[p^C(S)\right]^2$$

Profits from supplying the product in both languages are given by:

[11] As in the adoption model, if λ is sufficiently close to 2 then a symmetric equilibrium in pure strategies does not exist.

$$\pi^{SC}(\beta_s, \beta_c) = \alpha \left[p^C(C)\right]^2 + (1-\alpha)\left[p^S(S)\right]^2 - F$$

In equilibrium, firms must be indifferent between those options which are effectively used. In particular, the equilibrium candidates of β_s and β_c are given by the system of equations $\pi^S(\beta_s, \beta_c) = \pi^C(\beta_s, \beta_c) = \pi^{SC}(\beta_s, \beta_c)$. If we solve the system disregarding the condition $\beta_s + \beta_c \leq 1$, then the equilibrium candidates are given by:

$$\widehat{\beta}_s(z) = \begin{cases} 0, \text{ if } z \leq 1 - \frac{\lambda}{4} \\ \frac{1}{2} + \frac{2(z-1)}{\lambda}, \text{ if } 1 - \frac{\lambda}{4} \leq z \leq 1 + \frac{\lambda}{4} \\ 1, \text{ if } z \geq 1 + \frac{\lambda}{4} \end{cases}$$

$$\widehat{\beta}_c(z) = \begin{cases} 0, \text{ if } \frac{\alpha z}{1-\alpha} \leq 1 - \frac{\lambda}{4} \\ \frac{1}{2} + \frac{2\left(\frac{\alpha z}{1-\alpha} - 1\right)}{\lambda}, \text{ if } 1 - \frac{\lambda}{4} \leq \frac{\alpha z}{1-\alpha} \leq 1 + \frac{\lambda}{4} \\ 1, \text{ if } \frac{\alpha z}{1-\alpha} \geq 1 + \frac{\lambda}{4} \end{cases}$$

It is also the case that $\widehat{\beta}_s(z) + \widehat{\beta}_c(z)$ is sometimes higher than 1. Thus, we can divide the parameter space in two regions. Let us label Condition B :

$$\frac{1-\alpha}{\alpha}\left(1 - \frac{\lambda}{4}\right) > 1 + \frac{\lambda}{4}$$

i.e.,

$$\alpha < \frac{1}{2}\left(1 - \frac{\lambda}{4}\right)$$

Note that Condition A implies Condition B, but not viceversa. Also, Condition B is necessary and sufficient to obtain $\beta^e = 0$ in the adoption model.

If Condition B holds, then for all z such that $\widehat{\beta}_c(z) > 0, \widehat{\beta}_s(z) = 1$. Thus, in equilibrium $\beta_c^e = 0$ and $\beta_s^e = \widehat{\beta}_s(z)$. That is, if $z > 1 + \frac{\lambda}{4}$ then all varieties are exclusively supplied in S. As z falls below $1 + \frac{\lambda}{4}$ and increasing number of varieties is supplied in both languages.

If Condition B fails then both β_s^e and β_c^e can be strictly positive. If $z > 2(1-\alpha)$ then there will be no translation ($\beta_s^e + \beta_c^e = 1$) and $\beta_s^e = \widehat{\beta}_s(2-2\alpha)$, $\beta_c^e = \widehat{\beta}_c(2-2\alpha)$. In this case the equilibrium values coincide with those computed in the adoption model . If $z < 2(1-\alpha)$ then the number of translations is strictly positive, and $\beta_s^e = \widehat{\beta}_s(z)$, $\beta_c^e = \widehat{\beta}_c(z)$.

Once again, it is immediate to check that: (i) $\beta_s^e \geq \beta_c^e$, (ii) neither β_s^e nor β_c^e decrease with F, and (iii) both β_s^e and β_c^e may increase or decrease with λ.

We can now compare the equilibrium and the first best allocations. Since we have two endogenous variables we could in principle discuss the discrepancy between social and private incentives to supply goods in each language, and hence split the parameter space in four regions, depending on whether $(1-\beta_h^e)-(1-\beta_h^*)=\beta_h^*-\beta_h^e$ is positive or negative for each language h, $h=S,C$. However, it is much more convenient to define an index of linguistic diversity. In particular, let μ be the difference between the total number of varieties in C and the total number of varieties in S:[12]

$$\mu \equiv (1-\beta_s)-(1-\beta_c)=\beta_C-\beta_S$$

We can say that an equilibrium exhibits insufficient (excessive) linguistic diversity if $\mu^e < \mu^*$ ($\mu^e > \mu^*$). It turns out that that the two regions are simply defined by a threshold value, $\widehat{z}$. If Condition B holds then $\widehat{z}=1$, otherwise $\widehat{z}=2\left(1-\alpha\right)\left(1-\frac{\lambda}{4}\right)<1$.

This discussion is summarized in the next Proposition

Proposition 2 *The equilibrium of the translations model may exhibit insufficient or excessive linguistic diversity. More specifically, there exists a threshold value of $z, \widehat{z}$, such that if $z > \widehat{z}$ then $\mu^e \leq \mu^*$ (if $\mu^* > -1$ then $\mu^e \leq \mu^*$), and if $z < \widehat{z}$ then $\mu^e \geq \mu^*$ (if $\mu^e < 0$ then $\mu^e > \mu^*$).*

Figure 4 depicts μ^e and μ^* in the three possible scenarios, depending on whether or not Conditions A and B hold. In particular, Figure 4a correspond to the case that both conditions hold, Figure 4b to the case Condition B holds but Condition A fails, and Figure 4c to the case that both conditions fail.

The literature on product diversity under monopolistic competition has emphasized that markets may generate too little or too much variety. The last proposition apparently conveys a similar flavor of ambiguity. However, standard models of product variety provide little guidance about the set of circumstances under which market bias has one sign or the opposite. In contrast, in our model excessive or insufficient linguistic diversity are closely

[12] The main advantage of this index is that it preserves the linearity with respect to z. An obvious alternative would be to define the index as the proportion of varieties in C: $\mu=\frac{1-\beta_s}{1-\beta_s+1-\beta c}$. The main qualitative results are the same.

linked to observable variables. Thus, if we observe a relatively small fraction of varieties in the minority language then the model suggests that there is insufficient linguistic diversity. However, if the fraction of translations is relatively high then there is excessive linguistic diversity.[13]

Let us consider the following example. If Condition B holds then at $z = \widehat{z} = 1$ we have that $\beta_c^e = 0$ and $\beta_s^e = \frac{1}{2}$. Thus, all varieties are available in S, and one half are available in C. Hence, we need to worry about excessive linguistic variety only when the fraction of varieties available in the minority language is more than $\frac{1}{3}$, even though the relative weight of the minority linguistic community could be arbitrarily low. If condition B fails then at $z = \widehat{z} < 1$ we have that $\beta_c^e = 0$ and $\beta_s^e < \frac{1}{2}$. Hence, the number of translations is higher than $\frac{1}{2}$. Summarizing:

Remark 3 *The sign of the inefficiency is closely linked to the degree of linguistic diversity observed in equilibrium. In particular, excessive linguistic diversity requires that more than one half of all possible varieties are supplied in the minority language, independently of the relative size of the minority linguistic community.*

In order to gain some economic intuition about the result that the market may provide either insufficient or excessive linguistic diversity, let us consider the limiting case $\alpha = \frac{1}{2}$. In this case the size of both communities is the same and the equilibrium level of linguistic diversity is always efficient. In fact, $\beta_s^e = \beta_c^e \equiv \beta^e$ and $\beta_s^* = \beta_c^* \equiv \beta^*$. However, private incentives could be excessive or insufficient with respect to social incentives. It turns out that the level of translations in equilibrium and in the first best are given, respectively, by:

$$1 - 2\beta^e = \begin{cases} 1, \text{ if } z \leq 1 - \frac{\lambda}{4} \\ \frac{4(1-z)}{\lambda}, \text{ if } 1 - \frac{\lambda}{4} \leq z \leq 1 \\ 0, \text{ if } z \geq 1 \end{cases}$$

$$1 - 2\beta^* = \begin{cases} 1, \text{ if } z \leq 1 - \frac{\lambda}{2} \\ \frac{2(1-z)}{\lambda}, \text{ if } 1 - \frac{\lambda}{2} \leq z \leq 1 \\ 0, \text{ if } z \geq 1 \end{cases}$$

[13] The magnitude of the inefficiency is very asymmetric. If F is very high then the loss in total surplus associated to profit maximization may be quite large. However, if F is so small that profit maximization leads to excessive linguistic diversity, then the loss in total surplus is due to F being paid too many times and, hence, is likely to be very small.

Therefore, the number of translations in equilibrium, whenever is strictly positive, is socially excessive, except if the fixed cost is so low that both in equilibrium and in the first best all varieties are translated (See Figure 5).

It is important to note that, in this model, translations do not expand aggregate demand and they simply allow some consumers to access a better combination of linguistic and non-linguistic characteristics. In most spatial models with single-product firms and when aggregate demand effects are absent, there is a tendency towards excessive product variety. Private incentives to introduce a new variety are higher than social incentives because all customers of a new variety are stolen from rival firms.

In our model, suppliers are potentially multi-product firms and hence must be concerned about the origin of the potential consumers of the second version: some of them are stolen from rival firms but some others simply switch across different linguistic versions of the same variety (the so-called cannibalization effect). Thus, the strength of the business-stealing effect is directly proportional to the weakness of the cannibalization effect. If the number of translations is small then the cannibalization effect is relatively important, which moderates private incentives to translate. In this case the market outcome turns out to be close to the first best. However, as the number of translations increases the cannibalization effect gets weaker (the business stealing effect gets stronger) and private incentives to translate overgrow private incentives.

Summarizing, Proposition 2 is a combination of the underprovision of linguistic variety in the adoption model (Proposition 1) and the excessive private incentives to translate. If the fixed cost of translation is high then the number of translations is small and the forces behind Proposition 1 dominate. If the fixed cost is sufficiently low, then the dominant effect is the excessive private incentives to translate and we end up with excessive linguistic variety.[14]

[14]It may be important to note that firms' profits increase if the number of translations is reduced below the equilibrium level. The reason is that when each individual firm decides whether or not to translate it does not take into account the effect on other firms' profits (business-stealing). This result may explain why an important cartel of film distributors (with an effective monopoly power over translations) refuses to dub films into Catalan even when a large fraction of the cost is paid by the Catalan regional government.

4 Concluding remarks

In some cultural goods and media markets it seems prohibitively expensive to supply the same content in more than one language. The model presented in this paper predicts that, under laissez-faire, the level of linguistic diversity in these markets will be inefficiently low. From a positive viewpoint, this result may contribute to explain why consumption in minority languages is surprisingly low in regions with bilingual population (like the case of Catalonia discussed in the introduction). It could also rationalize some widespread concerns about the potential negative consequences of the expansion of English as a second language on the use and development of other languages. From a normative point of view, this result can justify on purely efficiency grounds certain public policies that aim at protecting minority languages.

However, the analysis also suggests that the development of new technologies that reduce the costs of translations is likely to improve the efficiency of market outcomes; and, in the limit, it could even result in an overprovision of goods in minority languages.

The main results of the paper are shown to be robust to changes in various specific assumptions of the base model: (i) whether or not members of the majority community are competent in the minority language, (ii) firms' ability to price discriminate between members of the various linguistic communities, (iii) firms' financing channels (charging a price versus advertising), and (iv) comparison of the equilibrium allocation with the first or the second best allocations. The latter point is particularly relevant for policy implications. In the context of the adoption model, for instance, the government can raise total welfare by using flat subsidies to firms supplying the good in the minority language.

The role of other assumptions is more difficult to assess. For instance, it was assumed, first, that all potential varieties are produced and, as a result, language choices do not affect aggregate sales, and, second, that consumer preferences over language and intrinsic characteristics are independent. These assumptions contribute decisively to the tractability and transparency of the model and they certainly have an impact on both the equilibrium and the first best allocations. However, they do not seem essential for the main qualitative results of the paper.

5 References

Caminal, R. and Ll. Granero (2008), Multi-product firms and product variety, mimeo.

Chen, Y. and M. Riordan (2007), Price and Variety in the Spokes Model, The Economic Journal, vol. 117, No. 522, 897-921.

Church, J. and I. King (1993), Bilingualism and network externalities, Canadian Journal of Economics, vol. 36, No.2, 337-45.

Dixit, A. and J. Stiglitz (1997), Monopolistic Competition and Optimum Product Diversity, American Economic Review 67 (3), 297-308.

Ginsburgh, V., I. Ortuño-Ortín, and S. Weber (2005), Disenfranchisement in Linguistically Diverse Societies: The Case of the European Union, Journal of the European Economic Association, Vol. 3, Issue 4, 946-65.

Ginsburgh, V., S. Weber, and S. Weyers (2007), Economics of literary translation: A simple theory and evidence, CEPR Discussion Paper No.6432.

Graddol, D. (2006), English next, published by the British Council. Available at: http:/www.britishcouncil.org/learning-research-english-next.pdf

John, A. and Kei-Mu Yi (2001), Language and Location, mimeo.

Lazear, E. (1999), Culture and Language, Journal of Political Economy, Vol. 107, Issue 6, 95-126.

Mankiw, G. and M. Whinston (1986), Free Entry and Social Inefficiency, Rand Journal of Economics 17 (1), 48-58.

Melitz, J. (2007a), The impact of English dominance on literature and welfare, Journal of Economic Behavior and Organization 64, 193-215.

Melitz, J. (2007b), Language and Foreign Trade, European Economic Review, vol. 52, Issue 4, 667-699.

Ortega, J. and T. Tangeras (2008), Unilingual Versus Bilingual Education System: A Political Economy Analysis, Journal of the European Economic Association, vol. 6, Issue 5, 1078-1108.

Peitz, M. and T. Valletti (2008), Content and advertising in the media: Pay-tv versus free-to-air, International Journal of Industrial Organization 26 (4), 949-965.

Salop, S. (1979), Monopolistic Competition with Outside Goods, The Bell Journal of Economics 10 (1), 141-156.

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

6 Appendix

6.1 The adoption model without price discrimination

Let p^s and p^c the prices set in equilibrium by a firm supplying a variety in S and in C, respectively. Hence, in contrast to the analysis of Section 2, firms cannot price discriminate according to linguistic preferences.

In those segments where consumers have access to one variety in each language, we denote by $x^C(C)$ and $x^C(S)$ the fraction of C-consumers that purchase the variety in C and in S, respectively. If both $x^C(C)$ and $x^C(S)$ belong to the interval $[0,2]$ then they are given by:

$$x^C(C) = 1 + \lambda + p^S - p^C$$

$$x^C(S) = 1 - \lambda + p^S - p^C$$

Let us first consider the case $\lambda < 1$. Then, both $x^C(C)$ and $x^C(S)$ belong to the interval $[0,2]$ if $1 - \lambda > p^S - p^C > \lambda - 1$. However, a firm can never find it optimal to supply the variety in C and set a price within this interval. More specifically, if firm i adopts S then total sales are given by:

$$q_i^S = \beta\left[1 + (1-2\alpha)\lambda + p^C - p_i\right] + (1-\beta)\left[1 + p^S - p_i\right]$$

Alternatively, firm i's sales from adopting C are given by:

$$q_i^C = \beta\left(1 + p^C - p_i\right) + (1-\beta)\left[1 - (1-2\alpha)\lambda + p^S - p_i\right]$$

For all p_i in the relevant range, $q_i^S > q_i^C$. That is, adopting C would imply a contraction of demand and hence it cannot be optimal. Therefore, if a firm adopts C it must be because it intends to set a sufficiently high price so that $x^C(S) = 0$. More specifically, firm i's profits from adopting C can be written as:

$$\pi_i^C = p_i\left\{\beta\left(1 + p^C - p_i\right) + (1-\beta)\,\alpha\left(1 + \lambda + p^S - p_i\right)\right\}$$

provided $1 + \lambda + p^S - p_i \leq 2$.

Evaluating the first order condition of an interior solution at $p_i = p^C$ we obtain the candidate of the "joint" reaction function of firms supplying their varieties in C:

$$p^C = \frac{\beta + (1-\beta)\,\alpha\left(1+\lambda+p^S\right)}{\beta + 2\,(1-\beta)\,\alpha} \tag{1}$$

Firm i's profits from adopting S can be written as:

$$\pi_i^S = p_i\left\{\beta\left[\alpha\left(1-\lambda+p^C-p_i\right)+2\,(1-\alpha)\right]+(1-\beta)\left(1+p^S-p_i\right)\right\}$$

provided $1-\lambda+p^C-p_i \geq 0$.

Evaluating the first order condition at $p_i = p^S$ we obtain the candidate of the "joint" reaction function of firms supplying their varieties in S:

$$p^S = \frac{1+\beta-\beta\alpha\left(1+\lambda-p^C\right)}{1-\beta+2\beta\alpha} \tag{2}$$

Thus, equations (1) and (2) can determine the equilibrium values of p^S and p^C for a given β.

The equilibrium value of β is given by the equal profits condition evaluated at equilibrium prices: $\pi^S(\beta) = \pi^C(\beta)$, which are given by:

$$\pi^S(\beta) = p^s\left\{\beta\left[\alpha\left(1-\lambda+p^C-p^S\right)+2\,(1-\alpha)\right]+(1-\beta)\right\}$$

$$\pi^C(\beta) = p^C\left\{\beta+(1-\beta)\,\alpha\left(1+\lambda+p^S-p^C\right)\right\}$$

We can now compute the parameter values for which $\beta^e \geq 0$. Note that $p^S(\beta=0) = 1$, and $p^C(\beta=0) = 1+\frac{\lambda}{2}$. Using these prices the inequality $\pi^S(\beta=0) \leq \pi^C(\beta=0)$ is equivalent to $\lambda \geq 2\left(\frac{1}{\sqrt{\alpha}}-1\right)$. Thus, in the absence of price discrimination, the set of parameters that can sustain an equilibrium with a positive fraction of C-products is smaller than in the case of price discrimination.

Equations (1) and (2) plus the equal profits condition determine the equilibrium values provided no firm has incentives to deviate from the proposed behavior. However, if λ is sufficiently high then firms producing a product in S find it optimal to deviate and set a higher price, p^d, $p^d = \frac{\beta(1-\alpha)}{1-\beta} - \frac{1+p^S}{2}$, and give up selling to C-consumers with access to a variety in C; i.e., $1-\lambda+p^C-p^d < 0$. In this case a symmetric equilibrium in pure

strategies does not exist. Note that if $\beta = 0$ there are no incentives to deviate. These incentives are present only if β is sufficiently high which involves a higher value of λ.

Numerical simulations can provide a good idea of the extent of the existence problem as well as the market bias against minority language in cases a symmetric equilibrium does exist. A selection of results are given in the following table (n.e. stands for "no symmetric equilibrium in pure strategies exists"):

α	λ	β^*	β^e
0.3	1.66	0.214	0.024
0.3	1.7	0.217	n.e.
0.35	1.4	0.267	0.054
0.35	1.5	0.275	n.e.
0.4	1.2	0.329	0.107
0.4	1.3	0.338	n.e
0.45	1	0.400	0.075
0.45	1.1	0.408	0.3
0.45	1.2	0.414	n.e.

6.2 Asymmetric bilingualism

Let us now consider the case that only C-consumers are bilingual, while S-consumers are monolingual. In other words, C-consumers incur a utility loss, λ, if they consume a product in S, but S-consumers experience an infinite utility loss if they consume the product in C.

Let us begin by computing the first best allocation for the case $\lambda < 1$. Among those consumers with access to two C-varieties only C-consumers can enjoy a positive surplus. Since they must consume the closest variety the maximum total surplus is $\alpha\left(R - \frac{1}{4}\right)$. The allocation of consumers with access to one variety in each language depends on their linguistic preferences. C-consumers are willing to consume the variety in S if their distance is lower than $1 - \lambda$, and S-consumers always consume the variety in S, independently of their location. Thus, the maximum total surplus is $\alpha\left(R - \frac{1}{4} - \frac{\lambda}{2} + \frac{\lambda^2}{4}\right) + (1-\alpha)\left(R - \frac{1}{2}\right)$. Finally, consumers with access to two varieties in S always choose the closest variety and the maximum total surplus is $R - \frac{1}{4} - \alpha\lambda$. Thus, total welfare can be written as:

$$W(\beta) = \beta^2\alpha\left(R - \frac{1}{4}\right) + 2\beta(1-\beta)\left[\alpha\left(R - \frac{1}{4} - \frac{\lambda}{2} + \frac{\lambda^2}{4}\right) + (1-\alpha)\left(R - \frac{1}{2}\right)\right] -$$

$$-(1-\beta)^2\left(R - \frac{1}{4} - \alpha\lambda\right)$$

Thus, the optimal fraction of varieties is:

$$\beta^* = \max\left\{0, \frac{\frac{\alpha}{2}(1+\lambda)^2 - \frac{1}{2}}{2(1-\alpha)\left(R - \frac{1}{4}\right) + \alpha\left(1+\lambda^2\right) - 1}\right\}$$

Note that $\beta^* > 0$ if and only if $\lambda > \frac{1}{\sqrt{\alpha}} - 1$.

In case $\lambda \geq 1$, the maximum total surplus for those consumers with access to one variety in each language is $R - \frac{1}{2}$ and the optimal fraction of varieties is:

$$\beta^* = \max\left\{0, \frac{2\alpha\lambda - \frac{1}{2}}{2(1-\alpha)R - \frac{3-\alpha}{2} + 2\alpha\lambda}\right\}$$

It can easily be checked that β^* is lower in cases where S-consumers are unable to enjoy a product in C.

In the market game competition for C-consumers is exactly like that in Section 2.3. However, firms supplying the good in C cannot compete for S-consumers. As a result, $p^S(S) = \frac{1+\beta}{1-\beta}$.

Profits are now given by:

$$\pi^S(\beta) = \alpha\left(1 - \frac{\lambda\beta}{2}\right)^2 + (1-\alpha)\frac{(1+\beta)^2}{1-\beta}$$

$$\pi^C(\beta) = \alpha\left(1 + \frac{\lambda(1-\beta)}{2}\right)^2$$

Again, in equilibrium $\pi^S(\beta^e) = \pi^C(\beta^e)$. In this case it is not feasible to compute β^e explicitly. But it can immediately be seen that $\beta^e > 0$ if and only if $\lambda > 2\left(\frac{1}{\sqrt{\alpha}} - 1\right)$. Thus, for those parameters that define the limits of this set, $\beta^* > 0$. In other words, asymmetric bilingualism simply reduces both the optimal fraction of varieties in C and the equilibrium fraction. But the qualitative properties of Proposition 1 still hold.

6.3 Advertising

Suppose consumers experience a utility loss which depends on the intensity of advertising, $\eta(A) = \frac{A^2}{2c}$. Thus, if consumers choose between two varieties in S then demand for variety i is given by:

$$x_i = 1 + \frac{A^2 - A_i^2}{2c}$$

where A_i and A are the advertising revenue of firm i and the average firm, respectively. Hence, firm i chooses A_i in order to maximize $\pi_i = A_i x_i$. In a symmetric equilibrium, $A_i = A = \sqrt{c}$. Thus, if $\beta = 0$ firms make profits $\pi^S(\beta = 0) = \sqrt{c}$.

If a firm j considers supplying the good in C then, for reasons discussed in Subsection 5.1, it aims at selling exclusively to C-consumers. In this case, the optimal level of advertising, A_j, maximizes:

$$\pi_j = A_j \alpha \left(1 + \lambda + \frac{c^2 - A_j^2}{2c}\right)$$

Thus, the maximum amount of profits is $\pi^C(\beta = 0) = \alpha\sqrt{c}\left(1 + \frac{2\lambda}{3}\right)^{\frac{3}{2}}$. The set of parameters that involve $\beta^e \geq 0$ is given by $\pi^S(\beta = 0) \leq \pi^S(\beta = 0)$. Note that this set is a subset of the set of parameters that support $\beta^e \geq 0$ in the case firms can extract consumer surplus through prices, but they cannot discriminate between members of different linguistic communities (Section 5.1). In other words, the advertising channel reinforces the underprovision of linguistic diversity.

6.4 The second best

Let us characterize the second best allocation; that is, the optimal β, denoted β^{**}, conditional on the allocation of consumers resulting from distortionary prices. In other words, in those segments where consumers have access to one variety in C and one in S, firms would set equilibrium prices computed in section 2.3. Thus, in these segments only those C-consumers located at a distance lower than $\frac{1}{2} - \frac{\lambda}{4}$ from the S-variety purchase this one and the rest purchase the C-variety. Similarly, only those S-consumers located at a distance lower than $\frac{1}{2} + \frac{\lambda}{4}$ from the S-variety purchase this one, and the rest purchase the C-variety. Total surplus obtained in these segments is given by $R - \frac{1}{4} - \frac{\lambda}{2} + \frac{3\lambda^2}{16}$.

Finally, we can write the total surplus obtained in a second best allocation as:

$$W^{SB}(\beta) = R - \frac{1}{4} - \beta^2(1-\alpha)\lambda + 2\beta(1-\beta)\left(-\frac{\lambda}{2} + \frac{3\lambda^2}{16}\right) - (1-\beta)^2\alpha\lambda$$

As a result, the second best level of βis :

$$\beta^{**} = \max\left\{0, \frac{1}{2} - \frac{4(1-2\alpha)}{3\lambda}\right\}$$

If we compare the above equation with those corresponding to the first best and equilibrium, we conclude that $\beta^* \geq \beta^{**} \geq \beta^e$, and the inequalities are strict except in the case that both variables take value zero. We can conclude that there is room for a simple policy intervention consisting on subsidies exclusively based on language choices.

6.5 The translation model for $\lambda > 1$

If $\lambda \geq 1$

$$W(\beta_c, \beta_s) = R - \frac{1}{4} - \alpha\left[\beta_s(1-\beta_s)\frac{1}{2} + \beta_s^2\lambda\right] - (1-\alpha)\left[\beta_c(1-\beta_c)\frac{1}{2} + \beta_c^2\lambda\right] - F(1-\beta_s-\beta_c)$$

and the optimal values of β_c^*, β_s^* are given by:

$$\beta_s^* = \begin{cases} 0, \text{ if } F \leq \frac{\alpha}{2} \\ \frac{\frac{F}{\alpha} - \frac{1}{2}}{2\lambda - 1}, \text{ if } \frac{\alpha}{2} \leq F \leq \alpha\left(2\lambda - \frac{1}{2}\right) \\ 1, \text{ if } F \geq \alpha\left(2\lambda - \frac{1}{2}\right) \end{cases}$$

$$\beta_c^* = \begin{cases} 0, \text{ if } F \leq \frac{1-\alpha}{2} \\ \frac{\frac{F}{1-\alpha} - \frac{1}{2}}{2\lambda - 1}, \text{ if } \frac{1-\alpha}{2} \leq F \leq (1-\alpha)\left(2\lambda - \frac{1}{2}\right) \\ 1, \text{ if } F \geq (1-\alpha)\left(2\lambda - \frac{1}{2}\right) \end{cases}$$

provided $\beta_c^* + \beta_s^* \leq 1$.

It can be easily checked that Proposition 2 also holds for $\lambda > 1$.

Figure 1

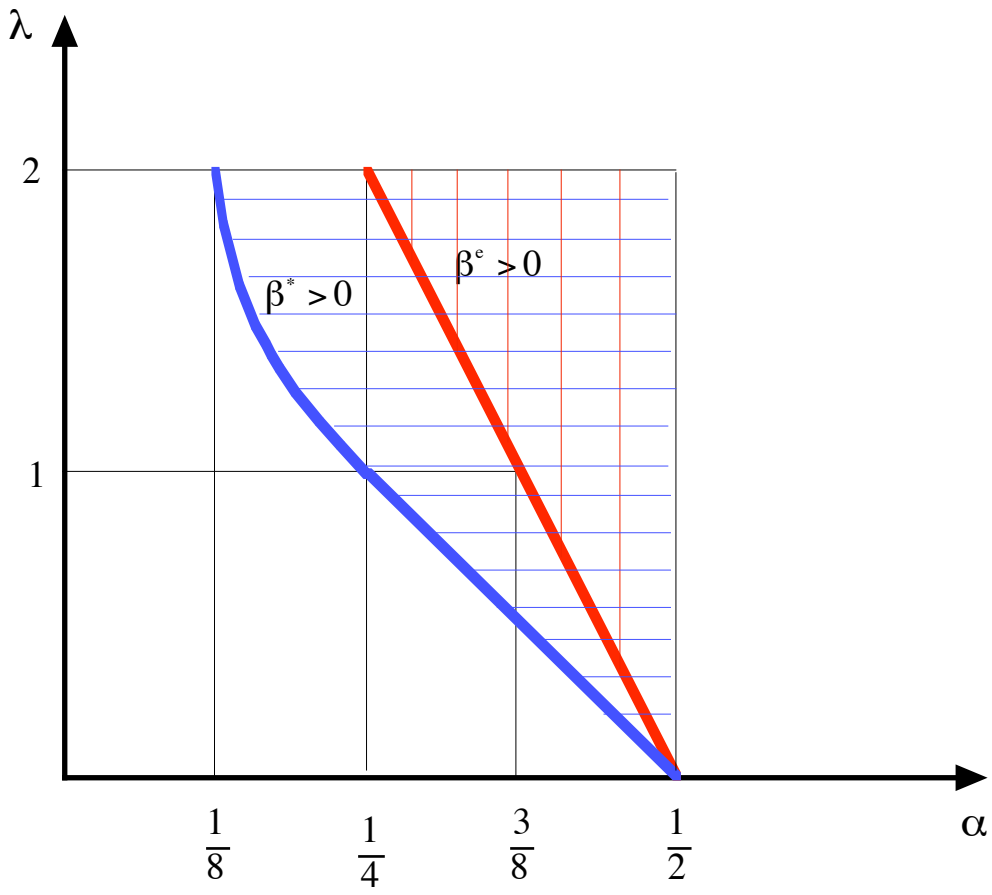


Figure 2

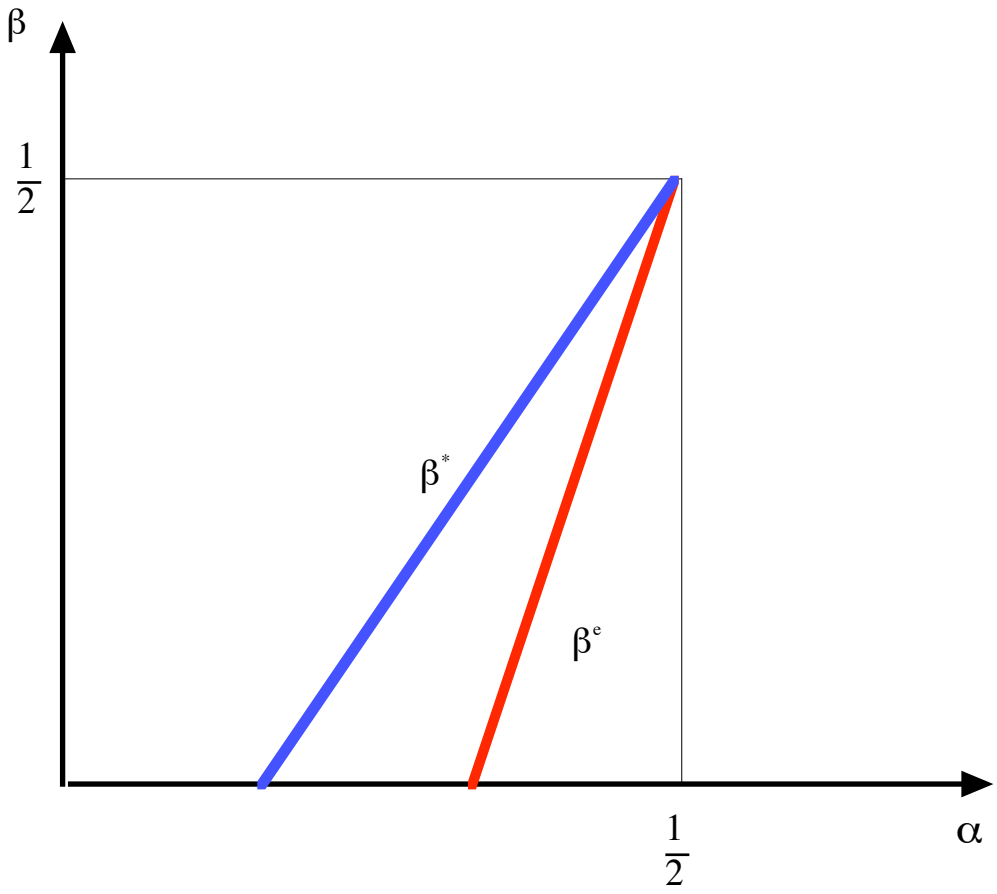


Figure 3

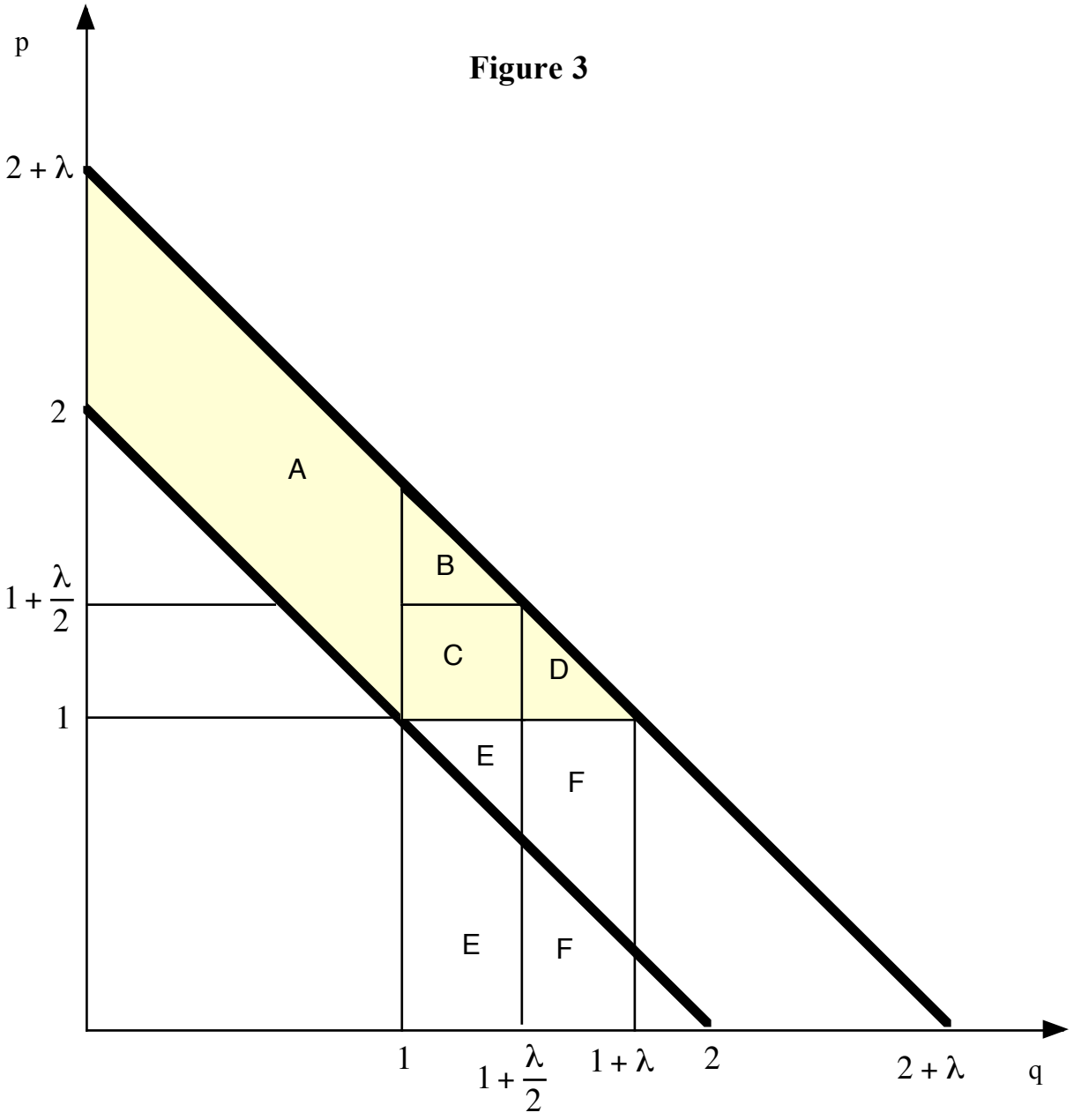


Figure 4a

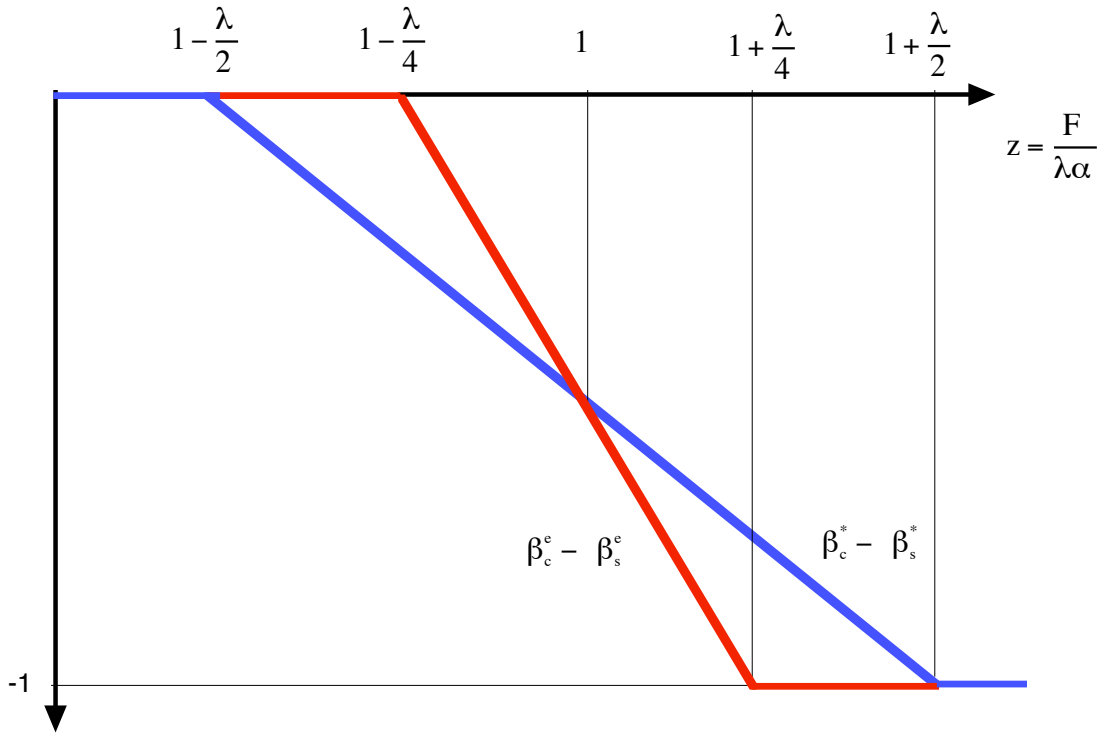

$1-\frac{\lambda}{2}$
$1-\frac{\lambda}{4}$
1
$1+\frac{\lambda}{4}$
$1+\frac{\lambda}{2}$
$z=\frac{F}{\lambda\alpha}$
$\beta_c^e-\beta_s^e$
$\beta_c^*-\beta_s^*$
-1

Figure 4b

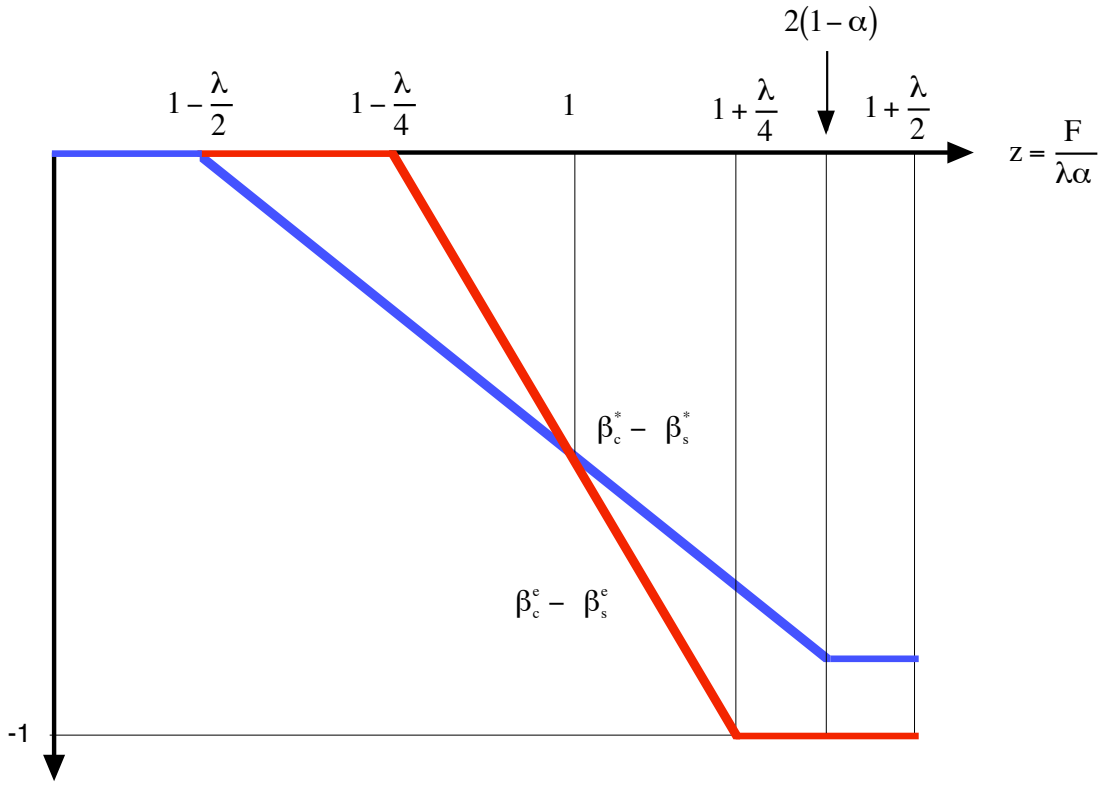

$2(1-\alpha)$
$1-\frac{\lambda}{2}$
$1-\frac{\lambda}{4}$
1
$1+\frac{\lambda}{4}$
$1+\frac{\lambda}{2}$
$z=\frac{F}{\lambda\alpha}$
$\beta_c^* - \beta_s^*$
$\beta_c^e - \beta_s^e$
-1

Figure 4c

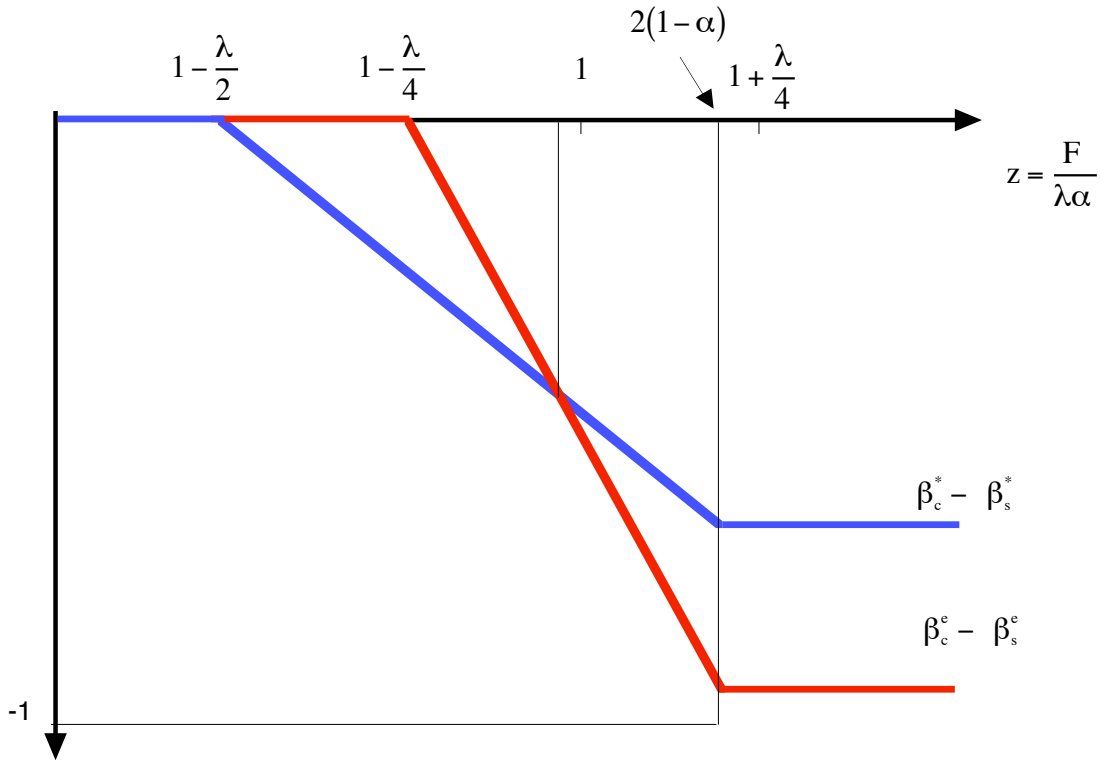
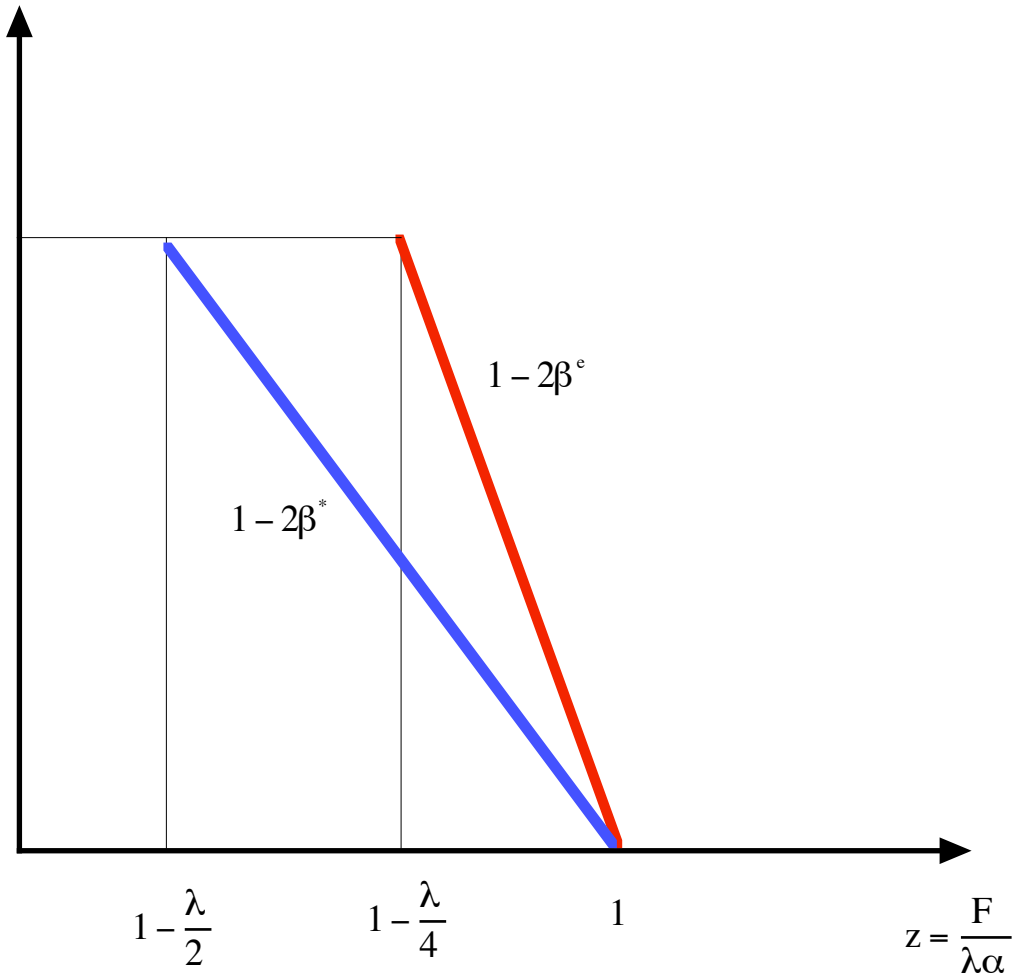

$1-\frac{\lambda}{2}$
$1-\frac{\lambda}{4}$
1
$2(1-\alpha)$
$1+\frac{\lambda}{4}$
$z=\frac{F}{\lambda\alpha}$
$\beta_c^* - \beta_s^*$
$\beta_c^e - \beta_s^e$
-1

Figure 5


$1-2\beta^{e}$
$1-2\beta^{*}$
$1-\frac{\lambda}{2}$
$1-\frac{\lambda}{4}$
1
$z=\frac{F}{\lambda\alpha}$