Chiswick, Barry R.

Working Paper

A Model of Destination Language Acquistion: Application to Male Immigrants in Canada

Working Paper, No. 149

Provided in Cooperation with:
George J. Stigler Center for the Study of the Economy and the State, The University of Chicago Booth School of Business

Suggested Citation: Chiswick, Barry R. (2000) : A Model of Destination Language Acquistion: Application to Male Immigrants in Canada, Working Paper, No. 149, The University of Chicago, Center for the Study of the Economy and the State, Chicago, IL

This Version is available at:
https://hdl.handle.net/10419/262551

Preliminary. For Comments Only. **Revised September 8, 2000**
Not for citation without authors' permission.

A MODEL OF DESTINATION LANGUAGE ACQUISTION: APPLICATION TO MALE IMMIGRANTS IN CANADA*

by

Barry R. Chiswick
Department of Economics
The University of Illinois at Chicago

and

Paul W. Miller
Department of Economics
The University of Western Australia

* We appreciate the comments received from the participants at the Conference on Immigration with an International Comparative Perspective, Vancouver, January 1999, European Society for Population Economics Annual Meeting, Turin, June 1999, Applications Workshop, Department of Economics, University of Chicago, May 1999, and the RIIM Seminar, Simon Fraser University, July 2000. This paper was written, in part, while Paul Miller was a Visiting Professor at the University of Illinois at Chicago, and while Barry Chiswick was John M. Olin Visiting Professor, Graduate School of Business, University of Chicago. Partial financial support for this project was provided through a grant from the Embassy of Canada, Washington, DC, from the George J. Stigler Center for the Study of the Economy and the State, Graduate School of Business, University of Chicago and from the Center for Excellence in Vancouver, Research on Immigration and Integration in the Metropolis (RIIM). Opinions expressed are those of the authors and should not be attributed to the sponsors.

September 2000

ABSTRACT

A MODEL OF DESTINATION LANGUAGE ACQUISTION: APPLICATION TO MALE IMMIGRANTS IN CANADA

Barry R. Chiswick
University of Illinois at Chicago

and

Paul W. Miller
University of Western Australia

This paper analyzes the determinants of language practice among adult male immigrants in Canada using the 1991 Census. A model is developed using human capital theory and an immigrant adjustment process. Language practice responds to exposure, efficiency and economic considerations. Use of English or French is greater the younger the age at migration, the longer the duration of residence, the higher the educational attainment, the further the country of origin from Canada, the linguistically closer is the mother tongue to English or French, among those who are not refugees, among those from a former British, French or American colony, and among those who live in an area where fewer people speak the respondent's mother tongue. (115 words)

Revised September 2000

A MODEL OF DESTINATION LANGUAGE ACQUISTION: APPLICATION TO MALE IMMIGRANTS IN CANADA

I. INTRODUCTION

It has been shown that language skills play an important role in determining the social and economic status of immigrants[1] Knowing the reasons for the language skills of immigrants is important for understanding the determinants of their economic well-being, as well as other aspects of economic, political, and social life in the destination. The identification of the groups "at risk" of lacking proficiency in an official language can provide a basis for the design of more effective public policies regarding immigration, language training, the labor market, and the social and political integration of immigrants.[2]

There has been two literatures, one in linguistics and one in the social sciences, regarding the determinates of destination language proficiency among immigrants. Much of this literature has focused on immigrants in the United States, and much of it uses ad hoc models. The linguistics literature tends to use small, selected (non-random) samples of the population, use simple statistical analyses, and to focus on one variable, age at immigration. The social science studies tend to use large random samples, from nationwide surveys or censuses, and to use multivariate statistical analyses, even if the focus is on one or two variables.

Perhaps the most fundamental variable in the analysis of immigrant acquisition of destination language skills that appears in nearly all studies, whether by linguists or social

---

[1] There is a large literature on the effects of destination language skills on labor market outcomes in the major immigrant receiving countries, including Australia, Canada, Germany, Israel, the Netherlands, the United States and the United Kingdom. For studies of Canada, see, for example, Abbott and Beach 1987, Chiswick and Miller 1988, 1992, 2000, Grenier 1987, and Shapiro and Stelcner 1997.

[2] That this is increasingly being recognized is well illustrated by the series of recent publications by Canadian Heritage on official languages (see Canadian Heritage, 1999). This is a multidisciplinary research series that provides valuable reference material on political, economic, institutional and community factors that are important in language studies. [Paul: Do we keep this footnote?]

scientists, is age at immigration (see, for example, studies by linguists, sociologists, demographers and economists in Asher and Garcia 1969, Bialystok 1997, Bongaerts, et.al., 1997, Chiswick 1991, 1998, Chiswick and Miller 1994a, 1998a, Espenshade and Fu 1997, Flege, et al., 1995, Johnson and Newport 1989, Long, 1990, Newport, 1990, Schumann 1975, Service and Clark 1993, C. Stevens 1999, and G. Stevens 1999). The basic finding is that language proficiency increases with a younger age at immigration, with a difference between the effects of age on the immigrant's own speaking skills and other forms of language skills (hearing, reading and writing). There does not appear to be any sharp age break, but rather that the impressive language acquisition ability observed among pre-teenage children is diminished at older ages of immigration.

The linguists focus on the biological or physical maturational developments that presumably influence the brain's functioning regarding learning language skills, particularly speaking and accents. The social scientists, on the other hand, expand on maturational factors by incorporating in their explanations for the effect of age at immigration various social and environmental factors, including schooling in the destination.

Another variable of primary interest to social scientists, but less so to linguists, has been duration of residence in the destination. This is usually described in terms of a greater "exposure" to the destination language or a greater "opportunity" (longer time) to learn the language (Chiswick 1991, 1998, Chiswick and Miller 1994a, 1998a, Grenier 1984, Loo, 1985, Jasso and Rosenweig 1990, C. Stevens 1999, G. Stevens 1992, 1994, 1999, Lindstrom and Massey 1994, Espenshade and Fu 1997). A curvilinear effect is typically found for the effect of duration. That is, destination language skills increase with length of time in the country, but the rate of increase diminishes with duration.

While the positive effect of duration on language skills is consistent with mere exposure to or opportunities to learn the language, the curvilinear effect suggests that the story is more complex. The curvilinear effect could arise if an asymptote is reached beyond which language skills either do not improve (perhaps because of complete proficiency), or improve only with considerable additional difficulty. The shape is also consistent with either an implicit or explicit investment model in which the rate of return

on an investment in language is greater if the investment is made earlier rather than later.[3] Yet, investments in language skills cannot be made instantaneously because beyond some point diminishing returns set in if greater investments are made per unit of time. As a result, large investments in language skills are made in the early period but investments diminish with duration, resulting in language proficiency increasing over time, but at a smaller rate the longer the duration of residence.

Various studies in the social sciences tend to include other variables, usually included as statistical controls rather than as variables of primary interest. These variables, which differ across studies, often include level of schooling, occupation, marital status, gender, region (state or province) of residence, urban or metropolitan area residence, school enrollment, and labor force attachment, among other variables.[4]

After deleting from the data observations that have or are expected to have a mother tongue that is the same as the destination's dominant language, country of birth or origin is typically treated in one of three ways. Some studies restrict the sample to immigrants from a particular country (see, for example, the Espinosa and Massey (1997) study of Mexican immigrants). Others do a "fixed effects analysis" by including dichotomous variables for country of birth in the statistical analysis. These studies find significant effects of country of birth, that is, when various personal characteristics are held constant, destination language proficiency still varies by birthplace (See, for example, Chiswick and Miller, 1998a). Still others do not include variables for country of birth, thereby creating omitted variables bias if included explanatory variables are correlated with birthplace and if birthplace for other reasons effects destination language proficiency. (See, for example, G. Stevens 1999). We are not aware of any studies that seek to analyze in depth the factors operating behind the birthplace effects.

This paper is an analysis of the determinants of the destination language proficiency of immigrants. In Section II it develops a model of investment in destination language skills. Among other implications, the model permits insights into the channels

[3] There are three reasons for this. New immigrants have lower earnings (lower opportunity cost of time devoted to the investment), they have a longer period to receive returns, and profitable investments are more profitable if made sooner rather than later.

[4] Gender differences in immigrant language proficiency are a primary focus of G. Stevens (1986) and Chiswick and Miller (1998a).

through which country of birth influences destination language proficiency. Section III discusses issues related to the data and empirical implementation. The model is tested using the 1991 Census of Canada.[5] Canada is a major immigrant receiving country that draws its immigrants from a wide range of countries of origin and linguistic origins. The microdata sample from the Canadian Census provides a wide range of relevant variables, as well as a very large sample size.

The empirical counterparts of the conceptual variables from Section II are developed in Section III for the Canadian data. This includes the construction of five variables based on country of birth and mother tongue that are central to the analysis. These variables, which have behavioral interpretations, are variables for the concentration of minority language speakers with whom the respondent shares a mother tongue, the physical distance (measured in miles) between the country of origin and the major gateway into Canada, the "linguistic distance" between the immigrant's mother tongue and English or French, whether the origin is a former British, French or American colony, and whether the immigrant is a refugee.

The empirical analysis for adult non-aged men is reported in Section IV. Multinominal logit analysis is employed because of the categorical nature of the dependent variable. Because of the unique role of French in Quebec, separate analyses are presented for all of Canada, "English" Canada and Quebec. The analyses are reported with birthplace dichotomous variables and with the five birthplace based variables replacing the country dichotomous variables. For ease of interpretation, predicted probabilities are also reported for these variables.

The paper closes (Section V) with a Summary and Conclusion. This section summarizes the paper, discusses policy implications and reports directors for future research.

II. A Model of Language Attainment

The model of language attainment is based on the assumption that language skills

[5] Other studies of the determinants of language skills among immigrants in Canada, include Richmond and Kalback (1980) and de Vries and Valee (1980), who cross tabulations, as well as Breton (1978a, 1978b), Veltman (1983) and Chiswick and Miller (1994a).

are a form of investment in human capital. Immigrants who are not already proficient in the dominant destination language(s) make optimal investments in dominant language acquisition. Investments in language skills may be made prior to or after immigration, and language skills affect the choice of destination (Chiswick and Miller 1994a).

(A) Language as Human Capital

Language skills satisfy the three requirements for human capital. First, these skills are productive in that they may increase earnings in the labor market or decrease costs of consumption (prices) by lowering the costs of communication with others.[6] Part of the productivity of language skills is that they increase the productivity of other forms of human capital. For example, the labor market productivity of a worker with professional skills compared to one who is just a laborer is greater for those proficient in the dominant language. That is, there is a complementarity between language skills and other forms of human capital (for English-speaking countries see, Chiswick and Miller, 1995, 1999, 2000, and for Israel see, Berman, et al., 2000). Second, language skills are also embodied in the person and, as with other forms of human capital, cannot be separated from the person. Finally, language skills are created at a "sacrifice," where the sacrifice is in the form of the time and out-of-pocket resources devoted to language acquisition by those making the investment and others who may be financing the investment in language skills. Thus, language skills are a form of human capital.

One can think of individual immigrants acting as if they are making optimal investments in destination language proficiency. They invest up to the point where the marginal rate of return to them from their investment is just equal to the marginal interest cost of the funds that they invest. A schematic representation is presented in Figure 1

[6] Studies using data from a variety of countries demonstrate the large positive effect of destination language skills on the earnings of immigrants. See, for example, for the United States (McManus, et al., 1983, Chiswick 1991, Chiswick and Miller 1995, 1999, Tainer 1988), Australia (Chiswick and Miller, 1995), Canada (Chiswick and Miller 1988, 1992, 2000), Israel (Berman, et al., 2000, Chiswick 1998), the Netherlands (Kee and van Ophem 1996), Germany (Dustmann 1994) and the United Kingdom (Shields and Wheatley-Price 1999). We are not aware of empirical research on the effect of destination language skills on the prices immigrants pay as consumers.

(Becker and Chiswick, 1966). The marginal rate of return from the investment is the demand curve for investment in language skills (D). This demand curve is higher the lower is the cost of obtaining language skills (whether foregone earnings or out of pocket costs), and the greater are the monetary and non-monetary benefits from investments in language skills. The cost of acquiring destination language skills are lower the greater is the immigrant's exposure to the destination language either pre- or post-immigration. The costs are also lower if the immigrant is more efficient in converting exposure into language acquisition. The benefits from destination language acquisition are greater the larger are the returns in the labor market through higher wages or greater employment (lower unemployment) among those with greater proficiency. The benefits are also greater the larger the non-monetary returns in the form of consumption benefits or from greater participation in the cultural, social or political life of the destination.

The demand curve for funds for investment in language skills is downward sloping because marginal rates of return on the investment decline with a higher level of investment. This decline occurs, in part, because more profitable investments are made sooner rather than later if investments are independent of each other. It also declines because foregone earnings (i.e., the opportunity cost of time devoted to the investment) increase with greater prior investments. Finally, there is a shortening in the remaining working life the more time that has been devoted to language training.

The marginal interest cost of funds for the investment is the supply curve for investment in language skills (S). This supply curve is lower or further to the right the greater the wealth and the greater the access to other resources for financing the investment (Figure 1). Those with greater wealth can obtain investment funds at a lower borrowing or self-financing cost than those who are poorer. The supply curve is upward rising or positively sloped because investors use cheaper sources of funds before using more expensive sources.

A rising marginal interest cost of funds (supply schedule) and a falling marginal rate of return (demand schedule) from additional investments implies that if investments take place there will be a stable equilibrium level of funds (I) devoted to investment in language skills (Figure 1). This translates into an optimal level of destination language proficiency for the individual. It is not possible to measure dollars invested in destination language proficiency or the schedules for the marginal interest cost of funds or the marginal rate of

return from the investment. It is, however, possible to measure destination language proficiency (LANG).

(B) Determinants of Language Proficiency

The above suggests that a reduced form equation could relate destination language proficiency among immigrants to factors that shift the supply and demand curves for investment, namely exposure, efficiency and economic factors affecting the demand for investment in language skills, and wealth or access to funds for investment in language skills affecting the supply of investment funds. Thus, the reduced form language proficiency equation can be written as:

$$LANG = f \text{ (Exposure, Efficiency, Economic Incentive, Wealth).}$$

(*i*) Exposure

Exposure to the destination language can occur pre-immigration or post-immigration. The model is trivial if immigrants come from an origin in which the primary language is the same as the primary language used in the destination. Yet, even for immigrants from countries in which the destination language is not the dominant language, there may be various levels of exposure in the origin prior to immigration. Thus, the characteristics of country of origin may matter.

An important country characteristic is the extent to which the destination language is used in the origin country. This is more likely if the country of origin is or had been a dependency of the destination. Thus, pre-immigration exposure to French is more likely among immigrants from the Ivory Coast and similarly exposure to English is more likely among immigrants from Ghana, in contrast to, say, immigrants from Angola. Thus, the colonial past of the origin is a relevant characteristic.

Post-immigration exposure to the destination language can be decomposed into two elements, time units of exposure and intensity of exposure per unit of time. Time units of exposure merely measures the number of years since immigrating to the destination. As indicated in the Introduction, because investments in language skills tend to be more profitable if made just after arrival, rather than delaying the investment, and

because improvements can be expected with practice, the effect of duration or years since migration will not be linear. A quadratic specification would allow for the improvement in language skills for each additional year in the destination to diminish with additional years since migration.

The intensity of exposure per unit of time in the destination can be related to "neighborhood" characteristics and to "family" characteristics. An immigrant who is not proficient in the dominant language can be more successful in avoiding exposure to and practice of the dominant language if the immigrant lives in an area in which many others use his or her origin language (Chiswick and Miller, 1996). It is not the use of the origin language per se that is the issue, but rather the ease of avoiding the dominant language. To the extent that there is a large community, community interactions and activities, newspapers, radio and even cable TV may be in the origin language. A larger linguistic community can support a linguistic enclave labor market. Working in a linguistic enclave would reduce exposure to the destination language. These effects can be measured by the proportion of the population of the area, regardless of nativity, that speaks the immigrant's origin language. Ethnic networks have been shown to be quite important in recent Canadian research (see, for example, Guindon and Poulin, 1998).

An intense environment in which the immigrant operates is the family. Language practice within the family will influence proficiency in the destination language. Marriage prior to immigration is more likely to be to a spouse with the same country of birth and linguistic background, and the origin language will be more likely to be used at home. The result would be less proficiency in the destination language. If marriage was to a person not proficient in the immigrant's origin language there would be greater usage of and proficiency in the destination language. Thus, the effect of marriage per se prior to immigration is likely to be a lower level of proficiency, but the effect of post-immigration marriage is ambiguous.

Children could have several, partially offsetting, effects on parental language skills. First, the greater ability of children to learn new languages and their intense exposure to the destination language in school tend to hasten children's linguistic adjustment (Long 1990 and Service and Clark 1993). By bringing home their improved destination language skills children serve as role models for and "teachers" of their parents.

Second, since children, whether native or foreign born, are more likely to acquire proficiency in the destination language than are their parents, they can serve as translators.[7] This role for children would detract from potential destination language proficiency among their parents. Since this role as translators is more important for consumption and home production activities than for labor market activities it would be expected to have a larger negative effect on the mother's than on the father's destination language proficiency.

Third, children also affect labor supply, particularly that of their mothers. To the extent that investments in destination language skills are made for their labor market benefits and to the extent that exposure to the dominant language in the labor market enhances proficiency, children would detract from parents' proficiency. Thus, children would be expected to have a less positive or more negative effect on their mother's than on their father's proficiency.

Finally, parents concerned with transmitting the cultural identity and practices of the country of origin to their children may retain in the home the language of the origin. To the extent that the parents use the origin language at home for the benefit of their children their own proficiency in the destination language is diminished.

This analysis of the effects of children on parental language skills suggests that the sign of the effect is ambiguous. It also suggests that the effects differ for the mother and the father – the effect of children on parental language skills would be less positive or more negative for the mothers than for the fathers.

(ii) Efficiency

Immigrants may differ in their efficiency in destination language acquisition. Efficiency refers to the extent of improvement in destination language skills per unit of exposure. One of the most important efficiency effects is age at migration. Younger people have a far greater capacity for learning a new language than do older individuals (Long 1990 and Service and Clark 1993). At least in part this seems to be a biological

[7] A New York Times article reported on the role of children as translators for their immigrant parents in the United States, and the negative effect that this has on their parents' acquiring English language skills. See, Hedges (2000).

process – the brain loses some of its flexibility in adapting to new languages as a person ages. As indicated in the Introduction, it is expected that destination language proficiency would decline with an older age at immigration, other variables being the same, but without any sharp breaks at particular ages.

School attainment may also have an efficiency effect. Those with higher levels of schooling may have a greater ability to learn, and this learning ability for school subjects may carry over to languages. Those with more schooling may also have learned more about their origin language, including its structure and grammar, and have a higher degree of proficiency in their origin language. A better or deeper understanding of one's own language may facilitate the learning of other languages. Moreover, if the destination language is an international language (such as English or French) those with higher levels of schooling in the origin may have had more exposure to the destination language in school prior to immigration. In some countries, for example, English or French is a mandatory subject of study at higher levels of schooling.[8] Finally, among immigrants who have not completed their schooling, post-migration schooling may be associated with higher levels of destination language proficiency because proficiency is required for entry into school and because the schooling in the destination language enhances proficiency.

The difficulty in learning a destination language depends in part on the person's origin language. It would be more difficult for a Chinese speaker to learn French than it is for a Spanish speaker to learn French because the differences between the languages are that much greater. That is, the "linguistic distance" between Chinese and French is greater than the distance between Spanish and French. The greater the linguistic difference between the destination and the origin language, the lower would be the efficiency of an immigrant for learning the destination language.

The approach that linguists take to language differences is based on their concern for the roots or evolution of languages. For example, the Cambridge Encyclopedia of Languages explains that: "The main metaphor that is used to explain the historical

[8] The finding among immigrants in Israel of a positive effect of schooling on proficiency in Hebrew (Chiswick 1998) suggests that exposure during schooling is not the sole mechanism explaining the relation between schooling and destination language skills in English speaking destinations.

relationships is that of the language family or family tree" (Crystal, 1987, p. 292, emphasis in the original. See also Grimes and Grimes, 1993). The Encyclopedia includes an extensive discussion of families of languages (Crystal, 1987, pp. 283-340). It has, however, only two brief paragraphs in a side bar box on "interlingual distance": "The structural closeness of languages to each other has often been thought to be an important factor in FLL (foreign language learning)… However, it is not possible to correlate linguistic difference and learning difficulty in any straightforward way, and even the basic task of quantifying linguistic difference proves to be highly complex, because of the many variables involved" (Crystal 1987, p. 371). As is shown in Chiswick and Miller (1998) and below, this is too pessimistic an interpretation.

Immigrants differ in their incentives for migrating. Those who move primarily for improved labor market opportunities for themselves are referred to as economic migrants. Those who move primarily to accompany or to join another family member, as is often the case for women and more so for children, are referred to as tied movers. Those who move because of a fear or perception of persecution or discrimination due to their race, ethnicity, political orientation or social class are referred to as refugees. Ideological migrants are those who move for nationalistic or political reasons to a destination that better satisfies these objectives, even if there is no fear or perception of near-term discrimination in the origin.

Refugees, tied movers and ideological migrants are less likely to be favorably self-selected than are economic migrants for labor market success in the destination since, by definition, this is of lesser importance in their migration decision. Since part of the adjustment is a linguistic adjustment, this implies that those who are not economic migrants are likely to have a lower ability in foreign language learning, that is, they would be less efficient in learning the destination language. Moreover, refugee status is often unanticipated and refugees may have had less time to plan for or prepare for the move than economic migrants. Thus, refugees, tied movers and ideological migrants would be expected to be less proficient in the destination language than otherwise similar economic migrants.

(*iii*) Economic Incentives

The economic incentives for destination language proficiency depend on the increment in wages from becoming proficient and the expected duration of employment

or duration of stay in the destination. One might want to set up a simultaneous system of equations in which language skills are a function of expected increments in wages, and wages are a function of language skills. This is not feasible for econometric reasons, primarily the difficulty in developing identifying variables. To some extent level of schooling may reflect a wage increment incentive effect since the economic returns to destination language proficiency increase with the level of schooling (Chiswick and Miller 1995, 1999, 2000 for analyses for several countries).

Another relevant economic incentive variable would be the likelihood of return migration, that is, the expected future duration in the destination labor market. The geographic distance of the country of origin from the destination is relevant here. The greater cost of migrating due to a greater distance has two inter-related effects. First, other things the same, immigrants will tend to be more favorably self-selected, and hence of a higher level of ability (efficiency), the greater is the geographic distance (Chiswick 1999). Second, distance is associated with a lesser expectation of return migration and hence a greater incentive to invest in destination specific skills, including language skills.

There are also "consumption benefits" defined broadly, for an immigrant obtaining destination language skills. These may be in the form of lower prices through more efficient search (or a broader market) for market goods and services, and greater participation in the social, political and cultural life of the destination country. If education increases the demand for social, political and cultural incorporation with the destination, then those with higher levels of education would have a greater economic incentive for obtaining destination language skills.

(iii) Wealth Effects

Finally, the supply side of the market for funds for investment in human capital needs to be considered. Variables that lower the interest cost of funds, that is, that shift the supply curve of funds for investment to the right, encourage greater investments in destination language proficiency. Access to the capital market is important. Greater wealth, which lowers the interest cost of funds, encourages investment in language skills and hence enhances language proficiency (Grenier and Vaillancourt, 1983). In the

absence of a more direct measure, greater wealth may be associated with a higher level of schooling.

(iv) The Overall Model

Thus, the model generates a conceptual equation (with hypothesized signs in parentheses):

> LANG = f [Age at migration (-), Years since migration (+), Education (+), Married prior to migration (-), Married after migration (?), Children (?), Linguistic Distance (-), Geographic Distance (+), Minority Language Concentration Index (-), Refugee (-), Colony (+)].

In the estimating equation there is, of course, a residual to account for purely random unmeasured individual characteristics that effect language proficiency. These include such factors as innate language ability and personality traits, (see, for example, Chastain 1975, Svanes 1987, Gardiner 1990).

III. The Data and Empirical Implementation

The primary data set used in this research is the 1991 Census of Canada, Public Use Microdata File (Individuals), 3 percent simple random sample of the population.[9] These data are supplemented from various sources. The construction of each of the variables used in the empirical analysis is described in detail in Appendix A. The Public Use Microdata File (Individuals) released from the 1991 Census contains information on the speaking knowledge of the official languages (English and French), knowledge of other languages, the language usually used at home, and mother tongue.[10]

---

[9] Since each observation has the same weight (33.3) the analyses are computed unweighted (or selfweighted).

[10] For an analysis of the language questions in the Census of Canada and recommendations for improvements, see Chiswick and Miller (1998b).

The language variables can be used to create a trichotomous dependent language variable (Appendix A). The first language state is the inability to conduct a conversation in either of the official languages of Canada, English and French (L1). The second language state is the ability to conduct a conversation in one or the other of the two official languages of Canada, but the person usually speaks a non-official language at home (L2). This language state may be considered a partial shift away from the language of the country of origin. The third language state is where the immigrant speaks English and/or French and usually speaks one or both of these official languages at home (L3). When an immigrant is from a country in which English and French are not spoken, and neither language is the respondent's mother tongue, language state L3 represents a complete shift from the language of the country of origin to an official language of Canada.

The statistical analysis is limited to adult foreign-born males 25 to 64 years of age. The analysis is limited to adult males who are not aged to focus on individuals likely to be fully attached labor market participants. The inclusion of females and post-age 65 males would necessitate expanding the analysis to include labor supply considerations which may be determined endogenously with destination language skills. (Chiswick and Miller, 1994b). Moreover, because of likely gender differences in the determinants of language proficiency, a simple dichotomous variable would be inadequate and separate equations would need to be computed by gender (Chiswick and Miller, 1998a, G. Stevens 1994). This would take the paper in directions that are beyond its scope.

Those born in identifiable English-speaking countries are excluded from the analysis, as they tend to arrive fully fluent in English. This includes the U.S., UK and those born in "Central America, Caribbean, Bermuda, and South America" with an English mother tongue.[11]

[11] French-speaking birthplaces (e.g., France and Belgium) are not separately identified in the Canadian Census. Few immigrants to Canada come from French-speaking countries.

Among adult male immigrants in Canada from non-English speaking countries, about 5 percent cannot speak either English or French (L1), 4 percent in Quebec and 5 percent in "English" Canada.[12] About 95 percent can speak English or French, of whom about half usually speak English or French at home and about half usually speak neither language at home. Thus, slightly less than half of these immigrants usually speak one of the official languages at home.

The Census provides basic information on the characteristics of the immigrants, as described in Appendix A. The variables for age, educational attainment, period of residence in Canada, marital status and province/metropolitan area are standard. Unfortunately, there are no data that permit the identification of whether the current marriage took place prior to migration or on the number of children for adult males.

As indicated in Section II, country of birth may play a key role in the empirical application of the human capital model of language skills. In Chiswick and Miller (1992), for example, eight country of birth dummy variables were included in the

[12] Language categories of male immigrants from non-English speaking countries, Age 25-64, (Percent):

Language State	Total Sample	English Canada[(a)]	Quebec
Speaks Neither English nor French (L1)	4.8	5.0	3.5
Speaks English and/or French AND:			
• Usually speaks a Non-Official Language at Home (L2)	46.6	46.4	48.5
• Usually speaks an Official Language at Home (L3)	48.7	48.6	48.0
Total[(b)]	100.0	100.0	100.0

(a) English Canada does not include the Atlantic Provinces.
(b) Columns may not sum to 100.0 due to rounding.

Source: 1991 Census of Canada, Public Use Microdata File (Individuals)

analysis of official language fluency in Canada in the 1981 Census, and the majority of these were highly significant and associated with sizable effects on the degree of language fluency. Yet dichotomous variables for country of birth may be thought of as a measure of what we do not know, that is, of the relevant unmeasured behavioral variables that are behind the country of birth effects.

The five variables discussed in Section II that reflect dimensions of country of birth are geographic distance, linguistic distance, the minority language concentration index, refugee status and former colony of an English or French speaking country. The incorporation of these variables for an analysis for Canada represents a major contribution of this paper.

The geographic distance variable reflects the cost in terms of money and time of moving from the origin to the destination, as well as the cost of a return migration.[13] It is measured as the distance, in thousands of miles, between the major city in the country of origin and the closest major Gateway City into Canada, Toronto or Vancouver. In analyses for Quebec province, Montreal is used as the Gateway City.

A second innovation in this study of Canada is a direct examination of the impact of "linguistic distance" on official language fluency. The greater the linguistic distance of the origin language from the destination language the poorer would be the proficiency in the destination language. Because of the multifaceted nature of languages, linguists have not been able to develop a measure of linguistic distance (Crystal, 1987, p. 292, 371). "Family trees" for languages reflects linguist's perceptions or hypotheses as to the origins of languages; not how difficult it is for the speaker of one language to learn another.

Chiswick and Miller (1998) developed an index of "linguistic distance" based on the degree of difficulty that Americans who are native English speakers have learning foreign languages (Appendix A). It is developed from a set of language learning scores (LS).[14] A low value of the score is indicative of a high degree of difficulty (e.g.,

[13] A similar variable has been used in studies of English language fluency in the U.S. by Espenshade and Fu (1997) and Chiswick and Miller (1998) to capture the effects associated with propensities for initial migration (self-selection) and return migration that vary by distance of the country of origin.

[14] These language scores are from Hart-Gonzales and Lindemann (1993).

Cantonese LS = 1.25) and a high value is indicative of a low degree of difficulty (e.g., Dutch LS = 2.75). Symmetry is assumed; if it is difficult for A-speakers to learn language B, then it is difficult for B-speakers to learn language A. In this study of language in Canada it is assumed that the linguistic distance (LD) index developed for English in the U.S. can also be applied to English in Canada. In the empirical application, linguistic distance is measured as the reciprocal of the language score, that is, LD = 1/LS. Thus, a higher value for LD means a greater distance between English and the origin language.

Some modification to the approach taken for the study of language in the U.S. is necessary here because of the dual-language nature of Canada. We are not aware of any comparable linguistic distance index for French. Three alternatives are pursued. First, the effect of the LD measure is allowed to vary between Quebec and English Canada. Second, a separate measure for countries speaking a Romance language is used to capture the linguistic distance from French. This is a Romance/Non-Romance language dichotomous variable used in the analysis for Quebec. Third, separate analyses are conducted for Quebec and English Canada.

A third innovation and constructed variable is the minority language concentration index. This is measured as the percentage of the population (native born and foreign born, male and female) aged 18 to 64 in the respondent's region of residence that reports the same mother tongue as the respondent.[15] The region is defined as the person's Census Metropolitan Area (CMA) or the balance of the province for those not living in a CMA.

Finally, two dichotomous variables are created that describe different characteristics of the country of origin. The precision of these variables is restricted by the limited detail on country of origin provided in the Canadian Public Use Sample. "Refugees" are identified as those reporting Vietnam or the USSR as their country of birth. Those born in former "colonies" of the U.S., UK or France are identified through having been born in South Asia, Vietnam, Africa, Hong Kong and the Philippines.

[15] The minority language concentration index has a substantial positive skewness. It cannot be entered in a logarithmic transformation because of the zero values. Tests regarding more complex specifications give essentially the same results as those obtained from the simple specification used in this analysis.

Table 1 lists means and standard deviations for the variables in the study of language. These data are for the 3 percent sample of foreign-born males from non-English speaking countries aged 25 to 64 years who report valid information on each of the variables used in the analysis. Data are reported for the total sample, and separately for Quebec and English Canada (other than for the Atlantic Provinces). For this group the mean age is 44 years for Canada. On average, almost one-half of these years have been spent in Canada, the mean duration of residence being 20 years. The mean educational attainment is 11.6 years. Quebec accounts for 17 percent of the sample, the Prairie Provinces for 13 percent, British Columbia for 15 percent, and Ontario for 54 percent. The major birthplace regions of the immigrants are Other Europe (24 percent), Italy (14 percent), South Asia (8.3 percent), and Africa, Germany and Portugal (6 percent for each).

There are few differences between the immigrant populations of Quebec and English Canada, other than for the birthplace distributions. There are relatively fewer immigrants from Germany, Poland, and the countries/regions of Asia (other than Vietnam), and relatively more immigrants from Italy, Africa, and Central and South America in Quebec than in English Canada. Most striking is the greater propensity for immigrants from Romance language countries and former French colonies to reside in Quebec, presumably because of the smaller linguistic distance (Chiswick and Miller, 1994).

IV. Empirical Analysis

The dependent variable in this analysis, LANG, has three categories, L1 for a person who does not speak English or French, L2 for a person who speaks English or French but usually does not do so at home, and L3 for a person who speaks English or French and usually does so at home.

Two statistical techniques can be applied, multinomial logit analysis or ordered logit analysis. Ordered logit analysis is the more restrictive of the two techniques. Ordered logit treats the three categories as ordered; going from a lower to a higher level of proficiency in one of the official languages. The three categories in the language variable are not strictly ordered in principle, although in practice that appears to be the situation. However, ordered logit assumes that the proportional odds of going from

category 1 to category 2 is the same as from going from category 2 to category 3. The hypothesis that the ordered odds are the same is not consistent with the data.[16] For these reasons the more flexible but somewhat more complex multinomial logit technique is employed and the logit coefficients are used to obtain predicted probabilities for various values of several explanatory variables.[17]

The analysis of the results of a multinomial logit model can focus on the estimated coefficients (which inform on the impact of variables on the log-odds), the associated marginal effects on predicted probabilities, or on predicted distributions across the language categories. The discussion here first considers only the sign and statistical significance of the estimated coefficients. When the final model is being discussed, as well as when the birthplace coefficients are examined, predicted probabilities of being in each of the language states are presented for a range of characteristics to simplify the exposition.[18]

Three models of language practice are estimated for Canada as a whole and separately for English Canada and Quebec. The first contains only variables for region of residence and personal characteristics other than birthplace (Table 2). The second augments the first core specification with variables for birthplace (Table 3). The third specification

---

[16] For the total sample, the chi-squared test statistics for the test of the proportional odds assumption in the ordered logit models corresponding to those presented in Tables 3, 4 and 6 are 357.8, 617.6 and 644.2, respectively. These have an asymptotic chi-squared distribution with 10, 24 and 15 degrees of freedom, respectively, and so indicate that the proportional odds assumption is not appropriate with these data and models.

[17] Chiswick and Miller (1997) show that there is little basis for choosing between the non-nested ordered probability and multinomial logit models in the study of language skills.

[18] The estimated coefficients in the multinomial logit model give the partial effects of the explanatory variables on the log of the odds of being in the second (L2) or third (L3) language state relative to being unable to speak either English or French (L1). A positive coefficient for log (L2/L1) means that the explanatory variable increases the probability of being in L2 relative to being in L1. The relevant log odds for L3 relative to L2 is easily calculated from: log (L3/L2) = log (L3/L1) – log (L2/L1).

The partial derivatives of the probability of being in L3, L2 or L1 with respect to the explanatory variables may be computed. For dichotomous variables with large coefficients these can be poor approximations of the true partial effects because the partial derivative concept refers to infinitesimal changes rather than discrete changes. For this reason predicted probabilities are preferred.

replaces the birthplace variables with variables for minority language concentration, linguistic distance, geographic distance (miles) between the country of origin and Canada, refugee status and a former British, French or American colony (Table 5).

It is apparent from Table 3 that age at migration, educational attainment and duration of residence are significant determinants of the measure of language practice. The coefficients in the equation for the log-odds of L3 to L1 are considerably larger in absolute value than in the equation for the log-odds of L2 to L1. This indicates that language shift to an official language increases with duration in Canada (with the largest increases coming in the early years) and with educational attainment.[19] Immigration at an older age, however, is associated with a lower probability of being in the third language state (speaks an official language and usually speaks an official language at home) than in the second language state (speaks an official language but usually speaks a non-official language at home), and a lower probability of being in the second language category than not speaking either of the two official languages.

Region of residence also has an influence on language outcomes among immigrants. Those living in Quebec and the Prairie Provinces have a greater use of English or French, compared to immigrants in Ontario, other things being the same. The reasons for these region effects are not clear, but these variables do mean that province-specific influences (fixed effects) are held constant. Those living in a Census Metropolitan Area (CMA) have a lower use of an official language than do immigrants living outside these major cities. This may be due to the limited concentration of foreign language speakers in rural areas. Marital status is, however, not a significant factor in the model.[20]

---

[19] The quadratic effect of duration on language practice follows from a human capital investment model. The lower wage (opportunity cost of time) of a new immigrant, a longer period to receive benefits and an incentive to make profitable investments sooner rather than later all encourage immigrants to make their investments in language skills as soon as they arrive. Because of eventual diminishing returns the more investments that are made in a given unit of time, these investments are not made instantaneously on arrival. And in particular with language capital, there is much learning by doing. Hence, the greatest investments are made in the early period and the intensity of investments diminish over time. This translates into proficiency increasing but at a decreasing rate with duration of residence.

[20] This may be a consequence of the discrepancies between the data available in the Individuals file (see Appendix A) and the ideal variable (see Chiswick and Miller (1992))

These broad patterns carry over in Table 2 to the separate analyses conducted for English Canada (column (ii)) and Quebec (column (iii)). The only notable difference in the results for the two distinct language regions of Canada is that the duration of residence effects have a pronounced non-linear pattern in English Canada, but are approximately linear in Quebec.

Table 3 augments the basic model of Table 2 with 14 dichotomous variables for birthplace. The reference group is immigrants from Italy. Inclusion of the birthplace dummy variables in the model has little impact on the magnitudes or levels of statistical significance of any of the variables in the original set of explanatory variables.[21] In this analysis, 20 of the 28 birthplace coefficients for Canada (total sample) are statistically significant at conventional levels, and as a group they are highly significant. They reveal that a number of birthplace groups, for example immigrants from China, have higher probabilities of not being able to speak an official language, while some other birthplace groups, such as Vietnam, have a higher probability of continuing to speak a non-official language at home. From the size of the estimated coefficients, compared for example to the coefficient on age at migration, and from the statistical significance of the coefficients, it is apparent that birthplace matters when one is trying to account for the distribution of language skills in Canada. This is further illustrated by the predicted distributions across the language groups presented in Table 4.

The predicted distributions presented in Table 4 are for two duration of residence groups: those who have resided in Canada for five years (left-hand side) and those who have resided in Canada for 15 years (right-hand side). The data for the more recent arrivals (predicted for a duration of 5 years in Canada) show that there is considerable variation across birthplace groups in the proportion unable to speak an official language. This ranges from around 30 percent for immigrants from China to 20 percent for immigrants from Portugal, to negligible proportions among immigrants from Germany, the Philippines and Africa. Among those immigrants who can conduct a conversation in an official language, there is also considerable variation in the propensity to speak a non-official language at

---

that distinguishes between immigrants married to an official language speaker and those who are married to a person with whom they share a mother tongue.

[21] The variation across countries in birthplace effects is reduced when these other variables are added to an equation with only birthplace variables.

home. This practice is relatively more frequent among immigrants from Italy, Poland, and the Asian countries.

The figures presented in the right-hand side columns of Table 4 (predicted for a duration of 15 years in Canada) illustrate the rapid growth in fluency with duration of residence, and also the shift away from the use of the language of the country of origin in most cases. Immigrants from Vietnam and China are notable exceptions to this general pattern, but very few of the Vietnamese have had a long duration of residence in Canada.

The information presented in Table 4 shows clearly that language practice and skill vary appreciably across the birthplace groups. The fundamental reasons behind this variation is the subject matter of the remainder of this section. Hence, Table 5 presents a model where the birthplace variables have been omitted, and measures for minority language concentration, linguistic distance, geographical distance of the country of origin from Canada, refugee origin and former colony of Britain, France or the United States are included in the estimating equation. As there is a one-to-one correspondence between the birthplace dichotomous variables and the measure of geographic distance, both cannot be included in the model at the same time. Moreover, the refugee and colony variables and, for some of the respondents, the minority language and linguistic distance measures have been computed using information on country of birth (see Appendix A).[22]

The five new variables have the expected signs and all are highly significant. The comparative performances of the models presented in Tables 3 and 5 can be assessed only informally, as the models are formally non-nested and yet the non-overlapping variables have been constructed using, in part, the same (birthplace) information. Compare the likelihood functions (P^2) and Pseudo R^2 in Tables 2, 3, and 5.[23] The addition of the five variables with behavioral interpretations (Table 5) substantially increases the fit over the equations without these variables (Table 2). The improvement in fit is even greater if these

---

[22] Approximately 32 percent of the sample have the linguistic distance measure constructed using information on birthplace. About 12 percent of the sample have the minority language concentration measure constructed using information on birthplace.

[23] In the analysis for Canada as a whole the Pseudo R^2 is 0.202 when there are no birthplace variables (Table 2), it increases to 0.245 when the five substantive variables

five variables are replaced by the set of birthplace dichotomous variables (Table 3). The five behavioral variables explain about two-thirds of the explanatory power attributable to the birthplace dichotomous variables. Thus, the birthplace variables provide a better "fit" which is most important for predictive purposes, but is clearly inferior for testing hypotheses and for understanding the underlying behavioral phenomenon.

The greater the geographic distance between the country of origin and Canada (and hence the greater the favorable selectivity and the less likely is return migration, *ceteris paribus*), the more likely it is that an immigrant will be able to conduct a conversation in an official language. And the more likely it is that the immigrant will use an official language at home.

An increase in the percentage of the population of the area (CMA or balance of province) that can speak the origin language of the immigrant is associated with an increase in the probability that the immigrant will not be able to speak an official language. It is also associated with an increase in the likelihood that an immigrant who can speak an official language will speak a non-official language at home. Thus, there is less ability and use of the official languages by immigrants the larger the proportion of the population in the area in which they reside who speak their origin language.

Exactly the same pattern of effects is associated with the linguistic distance measure. That is, where an immigrant's mother tongue is linguistically distant from English, and hence it is expected that it will be more difficult for the immigrant to learn English, there is less likelihood that the immigrant will be able to conduct a conversation in an official language, and if able to conduct a conversation in an official language, it is also less likely that an official language will usually be spoken at home.

This specification of the measure of linguistic distance in Table 5 does not take account of the dual-language nature of Canada. Recall that the linguistic scores used in the construction of the linguistic distance variable are defined with reference to English. There is a similarity of the signs and statistical significance of the results when separate estimates are obtained for English Canada and Quebec. The multinominal logit equations, however, for the total sample with interaction terms between Quebec and the linguistic distance variable suggests that the adverse effect on official language practice of a greater linguistic

---

are added to the equation (Table 5) but to only 0.267 when they are replaced by the

distance of the origin language from English is larger in Quebec than it is in English Canada.[24] This would be consistent with the hypothesis that for immigrants in Quebec with a large linguistic distance from English (*e.g.*, Asian and Middle Eastern languages) the reduced exposure to English reduces the language shift to English or French.[25]

For the sample of immigrants living in Quebec, a model was also estimated where a dichotomous variable for the Romance language countries replaced the linguistic distance measure. The rationale is that French is closer (linguistically) to other Romance languages than to the other languages. This variable was defined as unity for immigrants from Italy, Portugal, and Central and South America who do not have an English mother tongue. The empirical findings from this model are broadly consistent with the motivation behind the linguistic distance measure. Controlling for the refugee and colony variables, immigrants from Romance language countries are more likely to be in categories L1 and L2 than in L3.[26] That is, they are more likely to retain their origin language in the home even if they can speak an official language. The findings based on the Romance language variable are, however, markedly inferior to those presented for Quebec in Table 5. This suggests that the information on inter-language variation in the linguistic distance variable defined with reference to English is relevant to language acquisition in this province. This may arise, in part, because English and French are quite close linguistically, and in part because of the extensive use of English in Quebec.

Given the limitations on the information on country of birth in the Census of Canada, the refugee variable is defined as unity for immigrants from Vietnam and from what is now the former Soviet Union. Other variables the same, refugees are less likely to

---

birthplace dichotomous variables (Table 3).

[24] The coefficients on LD in the model with the interaction term for Log (L2/L1) are -1.328 ('t' = 3.91) and for Log (L3/L1) -6.887 ('t' = 18.66). The coefficients on the interaction term between Quebec and LD are Log (L2/L1) –2.959 ('t' = 4.05) and Log (L3/L1) –1.337 ('t' = 1.74), respectively.

[25] Chiswick and Miller (1994) show that immigrants in Canada, whether in English Canada or Quebec, tend to acquire proficiency in English rather than in French, but that in Quebec immigrants from Romance language countries are more likely to acquire French proficiency or become official language bilinguals.

speak an official language and if they can speak an official language they are less likely to report that they usually speak it at home (Table 5).

Immigrants from former colonies of Britain, France or the United States, on the other hand, are more likely to speak an official language. Among immigrants who can speak an official language, however, those from a former colony are more likely than other immigrants to speak their origin language in total Canada and English Canada. In Quebec, on the other hand, although the colony effect on speaking an official language is larger than in English Canada, there appears to be no effect of a colonial origin on whether an official language speaker usually speaks one of these languages at home.

The patterns of language practice associated with minority language concentration, linguistic distance, geographic distance, refugee status and former colony can be illustrated through the computation of predicted distributions similar to those presented in Table 4. These predictions are computed for two duration of residence categories, immigrants who have resided in Canada for 5 years and for 15 years. Predicted language probabilities are reported in Tables 6 through 8 for the five variables for Canada as a whole.

Table 6 reveals that residence in an area in which the origin language is spoken more intensely is associated with both a higher probability of being unable to speak an official language (higher L1), and mother tongue retention in the home even when the immigrant speaks an official language (higher ratio of L2 to L3). The effects of minority language concentration on mother tongue retention are smaller among those who have lived in Canada a longer period of time. Consider the effect, evaluated at the mean for the other variables, at a minority language concentration of 2 percent (close to the mean of 2.2 percent), for example, Arabic speakers in Montreal. After 5 years in Canada 8 percent of adult immigrant men cannot speak English or French and only 18 percent usually speak one or the other at home, in contrast to 6 percent and 22 percent, respectively for those living in areas where virtually no one (0.0 percent) speaks their origin language. By 15 years in Canada, at a 2 percent concentration ratio only 2 percent do not speak an official language, while 37 percent usually speak an official language at home. In contrast, for a high minority language concentration group, say 6.0 percent (the concentration ratio for Italian speakers in

---

[26] The coefficient (and t-ratio) in the Romance Language variable in the equation for log (L2/L1) is 0.033 ('t' = 0.11). In the equation for log (L3/L1) it is –1.400 ('t' = 4.36). Then log (L3/L2) = -1.433.

Windsor), at 15 years in Canada nearly 4 percent cannot speak English or French and only 28 percent usually speak it at home. Thus Table 6 demonstrates that the proportion not usually using an official language at home (L1 and L2) rises and the proportion usually speaking an official language at home (L3) declines with an increase in the minority language concentration ratio.

Table 7 shows that the geographic distance between the country of origin and Canada has a substantial impact on fluency rates, particularly among the more recent arrivals. The inability to speak English or French decreases with distance while the extent of usually speaking one of the official languages at home increases with distance. With increases in duration of residence in Canada, proficiency increases for all distances from the origin. After 5 years in Canada, the predicted proportion who cannot speak English or French declines with distance from 10.1 percent at 3,000 miles (approximately the distance from Bogota to Toronto) to 6.6 percent at 7,000 miles (approximately the distance from Ho Chi Min City to Vancouver). The proportion who usually speak English or French at home increases from 14.4 percent to 22.5 percent, respectively, for these same distances. At 15 years duration the proportion usually speaking English or French at home increases from 30.9 percent to 43.1 percent with the increase in the distance from 3,000 to 7,000 miles.

Cohorts with a longer duration of stay in Canada will contain fewer immigrants with an expectation of a return to their country of origin, in part because many of those who had a high expectation of returning did return, while others who stayed revised downward their expectation of returning prior to retirement. With a decline in the average expectation of returning there would be greater investment in Canada-specific skills, and a greater investment in Canada-specific skills would lower the propensity for return migration. This decline in expectations of return migration with duration will result in a smaller spread of predicted values across the "miles" measure in Table 7 at 15 years duration than at 5 years duration of residence in Canada.

Predicted distributions across language categories by linguistic score are reported in the top panel in Table 8. The three language scores listed, are 1.38, 2.00 and 2.75, the scores for Chinese, Polish, and Dutch, respectively. It is apparent that language score has a major impact on fluency rates and mother tongue retention. After 5 years in Canada, 13.1 percent of those with a language score of 1.38 (the value for Chinese speakers) but only 5.5 percent of those with a language score of 2.75 (the value for Dutch speakers) cannot speak English

or French, while only 6.1 percent of the former and 32.2 percent of the latter usually speak one of the official languages at home. By 15 years in Canada it is predicted that only 15.0 percent of Chinese speakers usually speak an official language at home and 80.6 percent can speak an official language but usually speak their origin language at home. For the Dutch-origin speakers these predicted proportions are quite different, 55.3 percent and 43.4 percent, respectively.

These data tell a simple and compelling story. The reason why some immigrants have not learned an official language of Canada, or usually speak their origin language rather than an official language at home even if they report they can speak the latter, is because it is much more difficult for them than for other immigrants to learn French or English. It follows that the attainment of a given level of fluency will require more investment if a person's mother tongue is, for example, Chinese than if the person's mother tongue is Polish, and the least investment among the groups studied if it is Dutch.

The predicted language practice for refugees and others is reported in the middle panel of Table 8. Refugees in Canada only a few years are far less likely to be able to speak an official language (20 percent cannot do so compared to 8 percent for others) and are far less likely to usually speak an official language at home (5 percent compared to 18 percent, respectively). Although official language skills and use increase with duration, even after 15 years a large gap persists on the basis of refugee origin.

On the other hand, immigrants from former British, French or American colonies (Table 8, bottom panel) are more proficient in an official language than others, but are more likely to usually speak their origin language at home. This holds at both 5 years and 15 years in Canada.

An interesting group is Vietnamese immigrants who are classified in this paper as both refugees and from a former colony. As refugees they would have a lower level of proficiency, but coming from a former colony would enhance their level of proficiency in English or French. Evaluated at the means of other variables, compared to immigrants who are not refugees or from a former colony, the Vietnamese are less likely to use English or French at home, even among those who can speak one of the official languages.[27]

[27] Predicted values of language categories by duration in Canada computed from Table 6:

5 Years			15 Years		
L1	L2	L3	L1	L2	L3

V. SUMMARY AND CONCLUSION

Language proficiency is an important aspect of immigrant adjustment. Knowing the language or languages that an immigrant can speak, and those that the immigrant usually speaks at home, and their level of proficiency in the destination language provides important information on the socioeconomic status of the immigrant, as well as the extent of integration into the social, political, cultural and economic life of the majority population in the host country.

This paper has been concerned with the determinants of destination language proficiency among immigrants that came from a different linguistic background than the primary or dominant languages of destination. In addition to the immigrant's own characteristics, it places particular emphasis on the characteristics of the origin as determinants of destination language proficiency. The empirical application is to immigrants in Canada, as reported in the 1991 Census.

A model of immigrant language acquisition is presented which is based on the supply and demand for funds for investment in language skills. This generates four fundamental variables: exposure to the official languages, efficiency in new language acquisition, economic incentives for acquiring destination language skills, and wealth to finance the investments in destination language training. Variables that represent empirical counterparts of these concepts are developed based on the data in the 1991 Census of Canada microdata files. Additional variables are added to the data file based on the respondent's country of birth and mother tongue. The analysis is conducted for adult (age 25 to 64) foreign-born men from non-English speaking countries of origin, using a 3 percent simple random sample of the population, with English and French treated as the official languages.

Due to the content of the language questions in the 1991 Census of Canada, the language categories are defined to represent three discrete levels of adjustment to

Vietnamese	11.29	83.63	5.09	3.81	83.65	12.53
REFUGEE = COLONY = 0	9.60	70.94	19.46	2.66	58.10	39.24

Canadian society. The first language category (L1) comprises the five percent of the adult male immigrant population that cannot conduct a conversation in either English or French. The second category (L2), approximately 47 percent of the adult male immigrant population, consists of those who can conduct a conversation in an official language but usually speak a non-official language at home. The third category (L3), representing about one-half of the adult male immigrant population, can speak English or French and usually speak an official language at home.

The analyses show that age at migration (an efficiency variable), educational attainment (reflecting efficiency, economic and wealth effects) and duration of residence in Canada (an exposure variable) are significant determinants of language practice. Immigration at an older age is associated with a lower probability of knowing an official language and, if it is known, usually speaking an official language at home. Shifting to an official language increases with duration in Canada and with educational attainment. Language practice also varies considerably across birthplace groups.

The analysis of the effect of birthplace is extended by the substitution of five variables that are constructed, in part, using information on birthplace and mother tongue. These variables are the geographic distance between the country of origin and Canada, the linguistic distance between the immigrant's mother tongue and English or French, the proportion of individuals living in the same region as the immigrant that speak his mother tongue, whether the person is a refugee, and whether the origin is a former British, French or American colony. Unlike dichotomous variables for country of birth, these variables have behavioral interpretations, and they provide for greater understanding of the factors affecting language practice among immigrants.

A greater geographic distance implies a greater positive selectivity in migration and a lower probability of return migration (efficiency and economic effects). The results show that the greater the geographic distance between the country of origin and Canada, the more likely it is that an immigrant will be able to conduct a conversation in an official language, and the more likely it is that the immigrant will use an official language at home.

A greater "linguistic distance" between the mother tongue and the official languages implies a higher cost of acquiring English/French language proficiency (efficiency effect). Empirically, a greater linguistic distance is associated with a lesser

use of the official languages. This holds whether the analysis uses the linguistic distance measure developed for English or the Romance language dichotomous variable used for Quebec.

An increase in the percentage of the population in the region in which an immigrant lives that uses the language that is the immigrant's mother tongue implies greater ease of avoiding use of the official languages (exposure effect). Empirically it is associated with an increase in the probability that the immigrant will not be able to speak an official language. It is also associated with an increase in the chances that an immigrant who can speak an official language will usually speak the origin language at home.

Refugees are found to be less likely to speak one of the official languages, and even among those who report they can do so, they are less likely to usually do so at home. This may be because of a less favorable selectivity, less planning for the move or a lesser commitment to the destination (efficiency effects) among refugees.

Immigrants from a former British, French or American colony are more likely to have been exposed to English or French in their country of origin (exposure effect). They are found to be more likely to be able to speak an official language, but among those who can speak an official language they are also more likely to usually use their origin language at home.

A model using the five country-based explanatory variables does not appear to fit the data statistically as well as the model based on a series of birthplace dichotomous variables. The difference in fit is relatively modest, however, and the behavioral models test a range of hypotheses and have greater intuitive appeal than does the model using a set of birthplace dummy variables.

The analysis has implications for public policy. The foreign born in Canada, or in any destination, will be more proficient in the destination language if the immigration policy focuses on younger immigrants, with higher levels of schooling, who are not refugees, but are from a geographically further origin, with exposure to the destination language in the country of origin. The adjustment is easier (less costly) if the origin and destination languages are linguistically closer. However, the adjustment is slower if the immigrants are concentrated in a small number of language groups or if they segregate themselves by language group, than if they are linguistically diverse.

Some of these characteristics can be explicitly incorporated into an immigration policy. These include age at migration, level of schooling, motive for migrating (e.g., economic migrants) and diverse origins. All of these characteristics inform on the likely difficulty (cost) of acquiring destination language skills among immigrants.

The analysis also suggests additional avenues for research. This study has been limited to men, but a comparable analysis for women would be fruitful, particularly if it incorporated labor supply and family composition effects on language practice. A feature of this analysis has been the limited data on country of origin provided in the microdata file from the Canadian Census. If more detailed data were to become available a richer analysis could be performed.

Additional research is needed on the mechanism through which linguistic concentrations retard destination language proficiency. Does this arise from labor market activities, household consumption activities, selective migration, or some combination of the three? If linguistic concentrations retard destination language proficiency, do they enhance retention of the origin language and, if so, which effect on language capital is greater? To what extent, if any, are labor market outcomes (e.g., earnings) effected by living in a linguistic concentration area, independent of the indirect effects on the labor market outcomes of these concentrations through the person's own language skills.[28]

Immigrant concentrations may arise on the basis of characteristics other than language, such as country of origin (national origin), ethnicity and religion. Do concentrations based on these other criteria have effects on destination language practice?

The measure of language proficiency used in this study was quite simple. Would other measures, particularly specific information on the degree of proficiency or test-based measures, for Canada and for other countries give similar results? Do attitudes toward the expectation of return migration and cultural assimilation effect destination language practice? Is there a relation between the degree of parental linguistic assimilation and the language proficiency, school performance and earnings of their Canadian-born children? A crude Romance language dichotomous variable is used as a measure of "linguistic distance" from French in the Quebec analysis. Can a more refined linguistic distance measure be developed for French, as was done for English?

[28] A preliminary answer for the United States is provided in Chiswick and Miller, 1999.

Further research is needed to decompose the effect of duration in Canada on language practice into its determinants. To what extent is it due to merely "learning by living," arising out of workplace experiences, or to formal English/French language classes?[29]

In conclusion, this paper has demonstrated that a systematic economic model can be developed and applied successfully to analyzing the language practice of immigrants in Canada. Variables reflecting exposure to English or French, efficiency in language acquisition, economic incentives for acquiring destination language skills and wealth are all relevant. In addition, it is shown that models with behavioral interpretations based on birthplace-related characteristics are statistically significant, intuitively more appealing and can be nearly as successful in statistically explaining language practice as a set of birthplace dichotomous variables. The analysis also suggests that this is a fruitful area for additional research.

[29] For research in the effects of intensive Hebrew language training programs on immigrants in Israel, see Beenstock 1996.

REFERENCES

Abbott, Michael G. and Charles M. Beach (1987). "Immigrant Earnings Differentials and Cohort Effects in Canada," Institute for Economic Research, Queen's University, Discussion paper No. 705.

Asher, James J., and Ramiro García (1969). "The Optimal Age to Learn a Foreign Language," *Modern Language Journal*, Vol. 53, pp. 1219-1227.

Becker, Gary S. and Barry R. Chiswick (1996). "Education and the Distribution of Earnings," *American Economic Review*, Vol. 56, No. 2, May, pp. 358-369.

Beenstock, Michael (1996). "The Acquisition of Language Skills by Immigrants: The Case of Hebrew in Israel," *International Migration*, Vol. 34, pp. 3-30.

Berman, Eli, Kevin Lang and Erez Siniver, (2000). "Language-Skill Complementarity: Returns to Immigrant Language Acquisition," Department of Economics, Boston University, xerox.

Bialystok, Ellen (1997). "The Structure of Age: In Search of Barriers to Second-Language Acquisition," *Second Language Research*, Vol. 13, pp. 116-137.

Bongaerts, Theo, Chantal van Summeren, Brigitte Planken and Erik Schils (1997). "Age and Ultimate Attainment in the Pronunciation of a Foreign Language," *Studies in Second Language Acquisition*, Vol. 19, pp. 447-465.

Breton, Albert (1978a). *Bilingualism: An Economic Approach*, Montreal: C.D. Howe Research Institute.

Breton, Albert (1978b). "Nationalism and Language Policies," *Canadian Journal of Economics*, Vol. 11, No. 4, November, pp. 656-668.

Canadian Heritage, (1999). *New Canadian Perspectives: Official Languages Research Publications, Publications Catalogue*, Canadian Heritage, Ottawa.

Charette, Michael F. and Ronald Meng (1998). "The Determinants of Literacy and Numeracy, and the Effect of Literacy and Numeracy on Labour Market Outcomes," *Canadian Journal of Economics*, Vol. 31, No. 3, August, pp. 495-517.

Chastain, Kenneth (1975). "Affective and Ability Factors in Second-Language Learning," *Language Learning*, Vol. 25, No. 1, June, pp. 153-162.

Chiswick, Barry R. (1991). "Speaking, Reading and Earnings Among Low-Skilled Immigrants," *Journal of Labor Economics*, Vol. 9, No. 2, April, pp. 149-70.

Chiswick, Barry R. (1998). "Hebrew Language Usage: Determinants and Effects on Earnings Among Immigrants in Israel," *Journal of Population Economics*, Vol. 11, No. 2, May, pp. 253-271.

Chiswick, Barry R. (1999). "Are Immigrants Favorably Self-Selected?" *American Economic Review*, Vol. 89, No. 2, May, pp. 181-85.

Chiswick, Barry R. and Paul W. Miller (1988). "Earnings in Canada: The Roles of Immigrant Generation, French Ethnicity and Language." In T. Paul Schultz (ed.) *Research in Population Economics*, Vol. 6, Greenwich, Connecticut: JAI Press Inc., pp. 183-228.

Chiswick, Barry R. and Paul W. Miller (1992). "Language in the Labor Market: The Immigrant Experience in Canada and the United States." In Barry R. Chiswick (ed.) *Immigration, Language and Ethnic Issues: Canada and the United States*, American Enterprise Institute, Washington D.C., pp. 229-296.

Chiswick, Barry R. and Paul W. Miller (1994a). "Language Choice Among Immigrants in a Multi-Lingual Destination," *Journal of Population Economics*, Vol. 7, No. 2, pp. 119-131.

Chiswick, Barry R. and Paul W. Miller (1994b). "Language and Labor Supply: The Role of Gender Among Immigrants in Australia," *Research on Economic Inequality*, Vol. 17, No. 2, pp. 153-189.

Chiswick, Barry R. and Paul W. Miller (1995). "The Endogeneity Between Language and Earnings: International Analyses," *Journal of Labor Economics*, Vol.13, No.2, April, pp. 246-288.

Chiswick, Barry R. and Paul W. Miller (1996). "Ethnic Networks and Language Proficiency Among Immigrants," *Journal of Population Economics*, Vol. 9, No. 1, pp.16-35.

Chiswick, Barry R. and Paul W. Miller (1997). "Modeling English Language Fluency Among Immigrants in the United States," Department of Economics, University of Illinois at Chicago, xerox.

Chiswick, Barry R. and Paul W. Miller (1998a). "English Language Fluency Among Immigrants in the United States," *Research in Labor Economics*, Vol.17, pp.151-200.

Chiswick, Barry R. and Paul W. Miller (1998b). "Census Language Questions in North America," *Journal of Economic and Social Measurement*, Vol. 25, pp. 73-95.

Chiswick, Barry R. and Paul W. Miller (1999). "Immigrant Earnings: Language Skills, Linguistic Concentrations and the Business Cycle," Stigler Center for the Study of the Economy and the State, Graduate School of Business, University of Chicago, Working Paper No. 152, June.

Chiswick, Barry R. and Paul W. Miller (2000). "The Complementarity of Language and Other Human Capital: Immigrant Earnings in Canada," Department of Economics, University of Illinois at Chicago, xerox.

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

De Vries, John and Frank G. Vallee (1980). *Language Use in Canada*, Ottawa: Statistics Canada.

Dustmann, Christian (1994). "Speaking Fluency, Writing Fluency and Earnings of Migrants," *Journal of Population Economics*, Vol. 7, No. 2, pp. 133-156.

Espenshade, Thomas J. and Haishan Fu (1997). "An Analysis of English Language Proficiency Among U.S. Immigrants," *American Sociological Review*, Vol. 62, April, pp. 288-305.

Espinosa, Kristin E. and Douglas S. Massey (1997). "Determinants of English Proficiency Among Mexican Migrants to the United States," *International Migration Review*, Vol. 31, No. 1, Spring, pp. 28-50.

Fitzpatrick, Gary L. and Marilyn J. Modlin (1986). *Direct-Line Distances: International Edition*, Metuchen, N.J.: The Scarecrow Press Inc.

Flege, James, Murray J. Munro and Ian R.A. MacKay (1995). "Effects of Age of Second-Language Learning on the Production of English Consonants," *Speech Communication*, Vol. 16, No. 1, pp. 1-26.

Gardiner, R. C. (1990). "Attitudes, Motivation, and Personality as Predictors of Success in Foreign Language Learning." In T.S. Parry and C.W. Stansfield (eds.), *Language Aptitude Reconsidered*, Englewood Cliffs, N.J.: Prentice Hall, pp. 179-221.

Green, Alan G. and David A. Green (1995). "Canadian Immigration Policy: The Effectiveness of the Point System and Other Instruments," *Canadian Journal of Economics*, Vol. 28, No. 4b, November, pp. 1006-1041.

Grenier, Gilles (1984). "Shifts to English as Usual Language by Americans of Spanish Mother Tongue," *Social Science Quarterly*, Vol. 65, No. 2, June, pp. 537-550.

Grenier, Gilles (1987). "Earnings by Language Group in Quebec in 1980 and Emigration from Quebec Between 1976 and 1981," *Canadian Journal of Economics*, Vol. 24, No. 4, November, pp. 774-791.

Grenier, Gilles and Francois Vaillancourt (1983). "An Economic Perspective on Learning a Second Language," *Journal of Multilingual and Multicultural Development*, Vol. 4, No. 6, pp.471-483.

Grimes, Joseph E. and Barbara F. Grimes (1993). *Ethnologue: Languages of the World*, Thirteenth Edition, Dallas: Summer Institute of Linguistics.

Guindon, Renae and P. Poulin (1998). *Francophiles in Canada. A Community of Interests*, Canadian Heritage, Ottawa.

Hart-Gonzalez, Lucinda and Stephanie Lindemann (1993). "Expected Achievement in Speaking Proficiency, 1993," School of Language Studies, Foreign Services Institute, Department of State, mimeo.

Hedges, Chris (2000). "Translating America for Parents and Family," *New York Times*, Monday, June 19, 2000, p. A19.

Jasso, Guillermina, and Mark R. Rosenzweig (1990). *The New Chosen People: Immigrants to the United States,* New York: Russell Sage.

Johnson, Jacqueline S. and Elissa L. Newport (1989). "Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language," *Cognitive Psychology*, Vol. 21, No. 1, pp. 60-99.

Kee, Peter and Hans von Ophem (1996). "Immigrant Wages in the Netherlands: The Role of Dutch Language Proficiency," in Saziye Gazioglu, ed., *Migrants in the European Labor Market*, Aderdeen, Scotland: J-Net Publishers, pp. 62-84.

Lindstrom, David P. and Douglas S. Massey (1994). "Selective Emigration, Cohort Quality, and Models of Immigrant Assimilation," *Social Science Research*, Vol. 23, No. 4, December, pp. 315-349.

Long, Michael H. (1990). "Maturational Constraints on Language Development," *Studies in Second Language Acquisition*, Vol. 12, No. 3, pp. 251-285.

McManus, Walter, William Gould and Finis Welch (1983). "Earnings of Hispanic Men: The Role of English Language Proficiency," *Journal of Labor Economics*, Vol. 1, No. 2, April, pp. 101-130.

Newport, Elissa L. (1990). "Maturational Constraints on Language Learning," *Cognitive Science*, Vol. 14, No. 1, pp. 57-77.

Richmond, Anthony H. and Warren E. Kalback (1980). *Factors in the Adjustment of Immigrants and Their Descendants*, Ottawa: Statistics Canada.

Schumann, John H. (1975). "Affective Factors and the Problem of Age in Second Language Acquisition," *Language Learning*, Vol. 25, No. 2, December, pp. 209-235.

Service, Elisabet and Fergus I.M. Clark (1993). "Differences Between Young and Older Adults in Learning a Foreign Language," *Journal of Memory and Language*, Vol. 32, pp. 608-623.

Shapiro, Daniel M., and Morton Stelcner (1997). "Language and Earnings in Quebec: Trends over Twenty Years, 1970-1990", *Canadian Public Policy/Analyse de Politiques*, Vol. 23, No. 2, pp. 115-140.

Shields, Michael A. and Stephen Wheatley-Price (1999). "The English Language Fluency and Occupational Success of Ethnic Minority Immigrant Men Living in English Metropolitan Areas," Paper presented at the CEPR Conference on Marginal Labour Markets in Metropolitan Areas, Dublin, October 10-12.

Statistics Canada (1994). *User Documentation for Public Use Microdata File on Individuals, 1991 Census of Canada*, Ottawa: Statistics Canada, Service Number 48-039E.

Stevens, Christine A. (1999). "Selection and Settlement of Citizens: English Language Proficiency Among Immigrant Groups in Australia," *Journal of Multilingual and Multicultural Development*, Vol. 20, No. 2, pp. 109-133.

Stevens, Gillian (1986). "Sex Differences in Language Shift in the United States," *Sociology and Social Research*, Vol. 71, No. 1, pp. 31-36.

Stevens, Gillian (1992). "The Social and Demographic Context of Language Use in the United States," *American Sociological Review*, Vol. 57, pp. 171-185.

Stevens, Gillian (1994). "The English Language Proficiency of Immigrants in the U.S." In Jeffrey Passel and Barry Edmonston (eds.), *Immigration and Ethnicity: The Integration of America's Newest Immigrants*, Washington, DC: Urban Institute, pp. 63-86.

Stevens, Gillian (1999). "Age at Immigration and Second Language Proficiency Among Foreign-Born Adults," *Language in Society*, Vol. 29, No. 4, December, pp. 555-578.

Svanes, Bjorg (1987). "Motivation and Cultural Distance in Second-Language Acquisition," *Language Learning*, Vol. 37, No. 3, September, pp. 341-359.

Tainer, Evelina (1988). "Engllish Language Proficiency and Earnings Among Foreign-Born Men," *Journal of Human Resources*, Vol. 23, No. 1, Winter, pp. 108-122.

Veltman, Calvin J. (1983). *Language Shift in the United States*, New York: Mouton.

APPENDIX A

Definitions of Variables

Definition of Population: Foreign-born men from non-English speaking countries, aged twenty-five to sixty-four. Non-permanent residents (*i.e.*, persons on a student authorization, employment authorization, Minister's permit or a refugee claimant) are excluded from the analysis as the 1991 Census Public Use Microdata File (PUMF) does not contain information on the year of entry into Canada for this group. A small number of persons for whom data were not available on questions used in the construction of variables, and those resident in the Yukon and Northwest Territories, are also excluded from the analysis. Other sample exclusions are noted in the definitions. For further information on the data set, see Statistics Canada (1994).

Language Practice (LANG): LANG is a trichotomous variable. The first category (L1) comprises individuals who cannot conduct a conversation in English or French. The second category (L2) comprises individuals who can conduct a conversation in English or French, but usually speak a non-official language at home. The third category (L3) comprises those who can conduct a conversation in English or French and usually use an official language at home.

Years of Education (EDUC): This variable records the total years of full-time education. It is constructed from the Census information on total years of schooling for respondents who do not possess a university qualification. For individuals who possess a university qualification, the following years of full-time equivalent schooling are added to the years of secondary schooling: Diploma below bachelor level (2.4 years); Bachelor's degree (three years for those reporting three or fewer years of university, four years for all others); Diploma above bachelor level (four years for those reporting four or fewer years of university, five years for all others); Degree in medicine, dentistry, etc. (seven years); Master's Degree (six years); earned doctorate (eight years).

Years Since Migration (YSM): The census information on year at arrival is presented in single years for some arrival cohorts, small intervals for some cohorts in the non-Atlantic provinces, and large intervals for the Atlantic provinces. A continuous measure was formed from this information by assigning midpoints to all arrival intervals, and subtracting this value from 1991. A quadratic specification is used.

Birthplace (BIRTH): The following countries or regions of birth are identified in the census file for immigrants living outside the Atlantic provinces (listed in order of numerical importance): United Kingdom; Other Europe; Central and South America and Caribbean; Italy; United States; Southern Asia (*e.g.*, Bangladesh, India, Pakistan); Federal Republic of Germany; Africa; Poland; Middle East and Western Asia (*e.g.*, Turkey, Iran, Lebanon, Saudi Arabia); Portugal; Other Eastern and South East Asia (*e.g.*, Japan, South Korea, Singapore, Thailand); People's Republic of China; Hong Kong; Philippines; Vietnam; USSR; Other. For immigrants resident in the Atlantic Provinces, the only birthplace categories distinguished are: United States; United Kingdom; Other Europe; and Other. Immigrants from the United Kingdom, the United States and those

from "Central America, Caribbean and Bermuda and South America" whose mother tongue is English are excluded from the analysis, given that study of language fluency is most appropriately focused on immigrants from non-English speaking backgrounds. In addition, the small number of immigrants from the residual "Other" birthplace region are excluded from the analysis, as a direct line distance can not be assigned to this group in the construction of the "MILES" variable (see below). Immigrants from Italy are used as the benchmark group.

Minority Language Concentration (CONC): Each respondent is assigned a measure equal to the percentage of the population aged eighteen to sixty-four in the region in which he lives that reports the same mother tongue as the respondent. The region is defined using information on residence in a Census Metropolitan Area and province of residence. The CMA is used, as is the balance of the province for those not living in a CMA. The non-official language groups of German, Netherlandic (*e.g.,* Dutch), Italian, Spanish, Portuguese, Polish, Ukrainian, Greek, Chinese, Austro-Asiatic (*e.g.,* Khmer, Vietnamese), Arabic, Punjabi, and Other Indo-Iranian (*e.g.,* Bengali, Hindi, Kurdish) are identified on the Census Public Use Microdata File (Individuals). Those reporting another language (other than English or French) are assigned the value zero on the assumption that the language incidence is trivial. Those reporting only English or French are assigned the mean value of the CONC variable for those from the same country of birth group.

Marital Status (MARRIED): Dichotomous variable that is set equal to one for individuals who are married (includes common-law partners) and is defined to equal zero for all other marital states.

Location: Two location variables are used in the study. The first records province of residence. This information was grouped as follows: Atlantic Provinces (Newfoundland, Nova Scotia, New Brunswick, Prince Edward Island), Quebec, Ontario, Prairie Provinces (Manitoba, Saskatchewan, Alberta), and British Columbia. The second locality variable records the size of the place of residence. Individuals residing in Census Metropolitan Areas (defined as having a population of at least 100,000 based on the 1986 Census) are distinguished from other individuals.

Age: Age is available in single years.

Linguistic Distance (LD): This variable is constructed from a measure of the difficulty of learning a foreign language for English-speaking Americans. It is based on a set of language scores (LS) measuring achievements in speaking proficiency by English-speaking Americans at the U.S. Department of State, School of Language Studies, reported by Hart-Gonzalez and Lindermann (1993). For the same number of weeks of instruction, a lower score (LS) represents less language facility, and, it is assumed, greater linguistic distance between English and the specific foreign language. For example, Italian is scored at 2.5 (in a range from one to three) and Arabic is scored at 1.5. This methodology assumes asymmetry across languages, that is, if a language is difficult for English-speaking Americans to learn, it is equally difficult for native speakers of that language to learn English.

The scores reported by Hart-Gonzalez and Lindermann (1993) are matched to the mother tongue codes used in the PUMF as follows:

Mother Tongue	Hart-Gonzalez & Lindermann language	Linguistic Score
English and/or French	- -	See Text
Aboriginal languages	- -	n.a.
German	German	2.25
Netherlandic	Dutch	2.75
Italian	Italian	2.50
Spanish	Spanish	2.25
Portuguese	Portuguese	2.50
Polish	Polish	2.00
Ukrainian	Russian	2.25
Greek	Greek	1.75
Chinese	Mandarin (1.5), Cantonese (1.25)	1.375
Austo-Asiatic languages	Cambodian	2.00
Arabic	Arabic	1.50
Punjabi	Hindi	1.75
Other Indo-Iranian	Hindi	1.75
Other	--	See Text

n.a. = not available. The small number of foreign born in this category are deleted from the sample.

In the construction of this variable, foreign-born persons who report English or French as the mother tongue, and persons in the "Other" home language category, are assigned a value of the linguistic score on the basis of their country of birth group. The values used for the birthplace categories are: Germany (German language, score of 2.25); Italy (Italian, 2.50); Portugal (Portuguese, 2.50); Poland (Polish, 2.00); USSR (Russian, 2.25), Hong Kong (Chinese, 1.375), China (Chinese, 1.375), Philippines (Tagalog, 2.00); Vietnam (Vietnamese, 2.00); Other Europe (mean score of countries of Other Europe computed for prime-age male immigrants in the 1990 U.S. Census, 2.23); Middle East and Western Asia (mean score of countries of Middle East and Western Asia computed for prime-age male immigrants in the 1990 U.S. Census, 1.89); Southern Asia (mean score of countries of Southern Asia computed for prime-age male immigrants in the 1990 U.S. Census, 1.91); Other Eastern and South East Asia (mean score for these regions computed for prime-age male immigrants in the 1990 U.S. Census, 1.36); Africa (mean score of countries of Africa computed for prime-age male immigrants in the U.S. Census, 2.11); Central America, Caribbean and Bermuda and South America (mean score for these regions computed for prime-age male immigrants in the 1990 U.S. Census, 2.25).

The method of computing scores for the broad birthplace regions assumes that the birthplace distributions and language backgrounds of immigrants in the US are the same as in Canada.

This assumption may not be strictly valid, but it is preferable to the alternatives of excluding this sizeable group from the study or assigning the mean of the language score for immigrants in Canada for whom valid scores could be computed.

The variable in the regression equations is linguistic distance, which is one divided by the linguistic score, that is LD = 1/LS.

In the analyses pooled across all regions of Canada, a shift variable for Quebec is also used in conjunction with the linguistic distance measure. That is, the specification is b_0LD + b_1QUEBEC*LD.

In addition, a variable for having a Romance language background (ROMANCE) is included in some specifications for Quebec.

Romance Language Background (ROMANCE): This is defined as being born in a Romance language country identified in the Census, namely Portugal, Italy, and Central and South America (excluding those of English mother tongue).

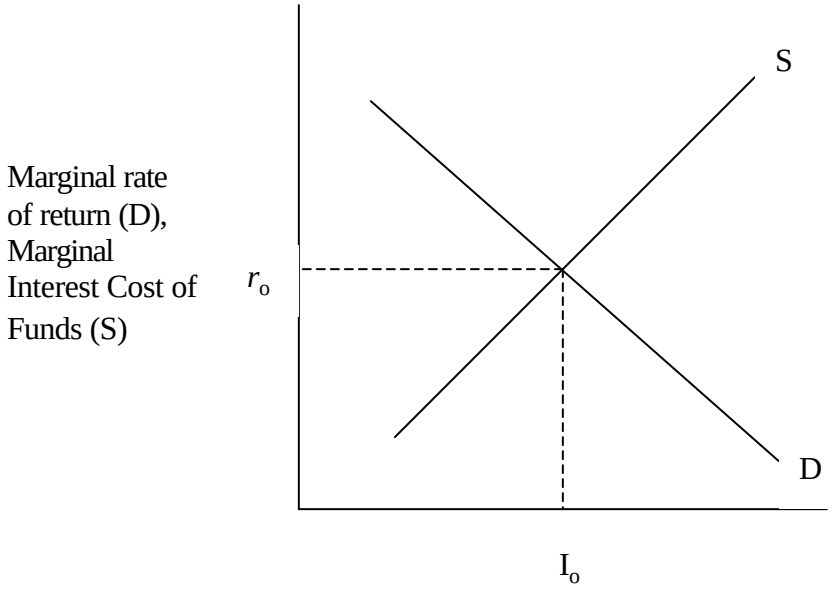
Direct-Line Distances (MILES): The miles between the major city in the immigrant's country of origin (or for broad regions a selected country within that region) and either Vancouver or Toronto, whichever was the smaller. For analyses limited to Quebec the variable records the distance to Montreal. The distances are from data in Fitzpatrick and Modlin's (1986) *Direct Line Distances, International Edition.*

Refugee (REFUGEE): Dichotomous variable equal to one for those born in Vietnam or the U.S.S.R., equal to zero for all other birthplaces.

Colony (COLONY): Dichotomous variable equal to one if born in a colony of Britain, the United States or France, that is, South Asia, Vietnam, Africa, Hong Kong, and the Philippines. Set equal to zero for all other birthplaces.

Figure 1

Schematic Representation of Supply and Demand for Funds
for Investment by Immigrants in Destination Language Capital



I_o : Optimal investment in destination language capital

r_o : Marginal rate of return at optimal investment

Table 1

Means and Standard Deviations of Variables, Male Immigrants from Non-English Speaking Countries, 1991 Census of Canada

Variable	Total Sample		English Canada[a]		Quebec	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Age	44.02	10.91	44.04	10.93	43.86	10.81
Age at Migration	24.18	11.46	24.00	11.51	25.30	11.06
Education Level	11.61	4.19	11.61	4.11	11.57	4.57
Period of Residence	19.84	12.64	20.05	12.73	18.57	11.96
Atlantic Provinces	0.006	0.08	(b)		(b)	
Quebec	0.170	0.38	0.000	0.00	1.000	0.00
Ontario	0.544	0.50	0.660	0.47	(b)	
Prairie Provinces	0.127	0.33	0.155	0.36	(b)	
British Columbia	0.153	0.36	0.185	0.39	(b)	
Lives in CMA	0.872	0.33	0.861	0.35	0.949	0.22
Married	0.815	0.39	0.819	0.38	0.794	0.40
Miles Between Canada & Origin[c]	4985	1468	5039	1443	4686	1726
Linguistic Distance	0.504	0.11	0.507	0.11	0.493	0.10
Minority Lang. Conc.	2.253	2.43	2.401	2.54	1.611	1.62
Refugee	0.048	0.215	0.050	0.218	0.042	0.200
Colony	0.253	0.434	0.265	0.442	0.199	0.400
Italy	0.135	0,34	0.126	0.33	0.181	0.39
Germany	0.062	0.24	0.070	0.26	0.022	0.15
Portugal	0.057	0.23	0.059	0.24	0.050	0.22
Poland	0.043	0.20	0.048	0.21	0.021	0.14
USSR	0.015	0.12	0.017	0.13	0.006	0.07
Other Europe	0.244	0.43	0.236	0.42	0.255	0.44
Middle East	0.048	0.21	0.039	0.19	0.091	0.29
Southern Asia	0.083	0.28	0.093	0.29	0.037	0.19
Hong Kong	0.044	0.20	0.051	0.22	0.007	0.09
China	0.050	0.22	0.057	0.23	0.022	0.15
Philippines	0.031	0.17	0.036	0.19	0.006	0.07
Vietnam	0.034	0.18	0.033	0.18	0.036	0.19
Other Asia	0.040	0.20	0.041	0.20	0.036	0.19
Africa	0.062	0.24	0.052	0.22	0.113	0.32
C. & S. America	0.053	0.22	0.040	0.20	0.116	0.32
Sample Size	32168		26,484		5,483	

(a) = Excludes the Atlantic Provinces. See text for explanation.
(b) = Variable not relevant.
(c) = Distance variable for Quebec define with reference to Montreal.

Source: 1991 Census of Canada, Public Use Microdata File (Individuals)

Table 2

Estimates of Logit Model of Language Practice, 25-64 Year Old Male Immigrants From non-English-Speaking Countries, 1991 Census of Canada

	Total Sample		English Canada[(a)]		Quebec	
	Log(L2/L1)	Log(L3/L1)	Log(L2/L1)	Log(L3/L1)	Log(L2/L1)	Log(L3/L1)
Constant	1.107	0.182	0.909	-0.097	3.750	3.801
	(5.46)	(0.85)	(4.26)	(0.43)	(3.32)	(3.35)
Age at Migration	-0.051	-0.091	-0.048	-0.091	-0.068	-0.099
	(16.56)	(27.42)	(14.57)	(25.30)	(8.02)	(11.02)
Educational	0.236	0.369	0.245	0.370	0.186	0.345
Attainment	(33.70)	(48.47)	(32.23)	(44.42)	(10.08)	(17.65)
Period of	0.119	0.219	0.120	0.228	0.118	0.196
Residence (PER)	(13.24)	(23.00)	(12.26)	(21.94)	(5.06)	(8.09)
$PER^2/100$	-0.087	-0.087	-0.090	-0.091	-0.088	-0.124
	(3.47)	(3.38)	(3.30)	(3.24)	(1.38)	(1.89)
Province (Ontario):						
Prairie Provinces	0.178	0.448	0.182	0.470	(b)	(b)
	(1.83)	(4.41)	(1.86)	(4.60)		
British Columbia	-0.179	-0.165	-0.187	-0.151	(b)	(b)
	(2.18)	(1.90)	(2.27)	(1.74)		
Quebec	0.585	0.799	(b)	(b)	(b)	(b)
	(6.58)	(8.63)				
Atlantic Provinces	0.598	1.285	(b)	(b)	(b)	(b)
	(0.59)	(1.26)				
Lives in CMA	-0.326	-1.154	-0.342	-1.051	-1.035	-2.795
	(2.58)	(8.98)	(2.65)	(8.00)	(1.01)	(2.74)
Married	0.128	-0.093	0.142	-0.050	-0.036	-0.303
	(1.55)	(1.09)	(1.60)	(0.53)	(0.16)	(1.29)
Sample Size	32,168		26,484		5,483	
c^2	11,041		9,815.6		1,356.1	
Pseudo R^2	0.2016		0.2160		0.1507	

(a) Excludes Atlantic Provinces: see text for explanation.
(b) Variable not relevant.
Note: L1 = Speaks neither English nor French; L2 = Speaks an Official language but usually speaks a non-official language at home; L3 = Speaks an Official language and usually speaks an official language at home.
Asymptotic 't' statistics in parentheses.

Source: 1991 Census of Canada, Public Use Microdata File (Individuals)

Table 3

Estimates of Logit Model of Language Practice, 25-64 Year Old Male Immigrants From non-English-Speaking Countries, Including Country of Birth, 1991 Census of Canada

	Total Sample		English Canada[a]		Quebec	
	Log(L2/L1)	Log(L3/L1)	Log(L2/L1)	Log(L3/L1)	Log(L2/L1)	Log(L3/L1)
Constant	0.940	-0.039	0.746	-0.256	3.707	3.628
	(3.89)	(0.15)	(2.91)	(0.94)	(3.14)	(3.05)
Age at Migration	-0.053	-0.102	-0.051	-0.103	-0.071	-0.108
	(16.17)	(28.19)	(14.43)	(26.13)	(7.80)	(11.26)
Educational	0.210	0.344	0.213	0.341	0.197	0.349
Attainment	(27.16)	(40.51)	(25.32)	(36.65)	(9.64)	(16.00)
Period of	0.141	0.234	0.143	0.243	0.146	0.219
Residence (PER)	(14.43)	(22.51)	(13.38)	(21.24)	(5.72)	(8.18)
$PER^2/100$	-0.124	-0.120	-0.130	-0.125	-0.125	-0.148
	(4.82)	(4.48)	(4.57)	(4.21)	(1.92)	(2.19)
Province (Ontario):						
Prairie Provinces	0.115	0.432	0.101	0.441	(b)	(b)
	(1.12)	(3.98)	(0.98)	(4.03)		
British Columbia	-0.130	0.052	-0.172	0.028	(b)	(b)
	(1.44)	(0.54)	(1.90)	(0.29)		
Quebec	0.499	0.558	(b)	(b)	(b)	(b)
	(5.38)	(5.73)				
Atlantic Provinces	0.164	0.338	(b)	(b)	(b)	(b)
	(0.16)	(0.33)				
Lives in CMA	-0.252	-0.847	-0.265	-0.754	-1.292	-2.895
	(1.93)	(6.31)	(1.98)	(5.48)	(1.25)	(2.81)
Married	0.260	0.083	0.277	0.131	0.140	-0.121
	(3.01)	(0.91)	(2.97)	(1.33)	(0.58)	(0.49)
Birthplace: (Italy):						
Germany[c]	12.915	14.135	13.035	14.187	11.955	13.244
	(0.09)	(0.10)	(0.09)	(0.10)	(0.02)	(0.02)
Portugal	-0.526	-0.404	-0.518	-0.472	-0.277	-0.001
	(4.35)	(3.08)	(3.95)	(3.32)	(0.84)	(0.00)
Poland	-0.006	-0.098	0.125	0.018	-0.796	-1.116
	(0.04)	(0.57)	(0.71)	(0.10)	(1.70)	(2.21)
USSR	0.776	0.761	0.845	0.698	12.815	13.738
	(2.31)	(2.19)	(2.49)	(1.97)	(0.01)	(0.01)
Other Europe	0.636	1.530	1.002	1.812	-0.515	0.499
	(4.47)	(10.56)	(5.93)	(10.57)	(1.81)	(1.71)
Middle East	0.914	1.179	0.885	1.172	1.043	1.048
	(4.93)	(6.08)	(4.27)	(5.39)	(2.46)	(2.38)
South Asia	0.862	0.821	0.911	0.826	1.792	1.988
	(5.45)	(4.96)	(5.45)	(4.72)	(2.19)	(2.39)
Hong Kong	1.029	-0.157	1.159	-0.046	-0.105	-1.543
	(4.87)	(0.70)	(5.24)	(0.19)	(0.13)	(1.66)
China	-0.952	-2.381	-0.802	-2.322	-1.965	-3.069
	(7.39)	(15.81)	(5.74)	(14.31)	(5.65)	(7.03)

Philippines	3.206	3.192	3.278	3.254	14.433	14.208
	(6.09)	(6.01)	(6.19)	(6.09)	(0.01)	(0.01)
Vietnam	-0.119	-1.571	-0.073	-1.539	0.240	-1.350
	(0.78)	(8.46)	(0.44)	(7.52)	(0.48)	(2.49)
Other Asia	0.352	0.092	0.568	0.377	-0.448	-1.494
	(2.04)	(0.50)	(2.91)	(1.82)	(1.18)	(3.45)
Africa	2.959	4.333	2.863	4.108	13.471	15.027
	(6.31)	(9.19)	(6.04)	(8.61)	(0.06)	(0.06)
C. & S. America	0.204	0.554	0.201	0.359	0.098	0.583
	(1.29)	(3.34)	(1.12)	(1.89)	(0.28)	(1.62)
Sample Size	32,168		26,484		5,483	
c^2	14,599		12,694		2048.3	
Pseudo R^2	0.2666		0.2794		0.2276	

(a) Excludes Atlantic Provinces: see text for explanation.
(b) Variable not relevant.
(c) There are no respondents from Germany in language category L1. The large estimated coefficients in the logit model bound the predicted probability to zero.

Note: L1 = Speaks neither English nor French; L2 = Speaks an Official language but usually speaks a non-official language at home; L3 = Speaks an Official language and usually speaks an official language at home.
Asymptotic 't' statistics in parentheses.

Source: 1991 Census of Canada, Public Use Microdata File (Individuals)

Table 4

Predicted Distributions across Language Categories by Birthplace and Duration of Residence, 1991 Census of Canada[(a)] (Percent)

	After 5 years in Canada			After 15 years in Canada		
Birthplace	L1	L2	L3	L1	L2	L3
Italy	12.38	73.83	13.79	3.43	65.32	31.25
Germany	0.00	61.24	38.76	0.00	38.16	61.84
Portugal	18.97	66.91	14.12	5.45	61.38	33.17
Poland	12.59	74.68	12.73	3.55	67.16	29.30
USSR	6.12	79.28	14.60	1.62	66.86	31.52
Other Europe	5.74	64.71	29.55	1.26	45.51	53.22
Middle East	5.13	76.29	18.58	1.28	60.80	37.92
South Asia	5.66	79.99	14.35	1.50	67.51	31.00
Hong Kong	5.37	89.53	5.11	1.61	85.84	12.54
China	29.36	67.61	3.03	10.87	79.96	9.16
Philippines	0.57	83.95	15.48	0.14	67.83	32.02
Vietnam	15.32	81.13	3.55	5.05	85.38	9.56
Other Asia	9.34	79.25	11.41	2.63	71.14	26.23
Africa	0.50	57.25	42.25	0.09	34.58	65.33
C. & S. America	9.75	71.33	18.92	2.49	58.07	39.44

Note: L1 = Speaks neither English nor French; L2 = Speaks an Official language but usually speaks a non-official language at home; L3 = Speaks an Official language and usually speaks an official language at home.

(a) = Predictions computed at means of all variables other than duration of residence. Row totals for each immigration period may not add to 100.0 due to rounding.

Source: Estimates in Table 3.

Table 5

Estimates of Logit Model of Language Practice, 25-64 Year Old Male Immigrants From non-English-Speaking Countries, Extended Model, 1991 Census of Canada

	Total Sample		English Canada[a]		Quebec	
	Log(L2/L1)	Log(L3/L1)	Log(L2/L1)	Log(L3/L1)	Log(L2/L1)	Log(L3/L1)
Constant	1.354	2.546	0.994	1.968	5.265	8.156
	(5.59)	(9.78)	(3.82)	(6.98)	(4.41)	(6.73)
Age at Migration	-0.050	-0.093	-0.047	-0.094	-0.064	-0.096
	(15.65)	(26.67)	(13.81)	(24.77)	(7.27)	(10.34)
Educational	0.234	0.388	0.240	0.389	0.201	0.365
Attainment	(31.94)	(48.10)	(30.14)	(44.01)	(10.30)	(17.54)
Period of	0.123	0.211	0.124	0.221	0.110	0.184
Residence (PER)	(13.38)	(21.69)	(12.41)	(20.69)	(4.61)	(7.36)
$PER^2/100$	-0.072	-0.064	-0.070	-0.066	-0.070	-0.087
	(2.86)	(2.48)	(2.56)	(2.33)	(1.09)	(1.32)
Province (Ontario):						
Prairie Provinces	0.128	0.488	0.105	0.496	(b)	(b)
	(1.27)	(4.60)	(1.03)	(4.64)		
British Columbia	0.062	0.337	0.049	0.344	(b)	(b)
	(0.66)	(3.41)	(0.52)	(3.45)		
Quebec	0.530	0.675	(b)	(b)	(b)	(b)
	(5.78)	(7.05)				
Atlantic Provinces	0.493	0.877	(b)	(b)	(b)	(b)
	(0.48)	(0.86)				
Lives in CMA	-0.115	-0.690	-0.112	-0.581	-1.031	-2.418
	(0.87)	(5.15)	(0.84)	(4.24)	(1.00)	(2.34)
Married	0.186	-0.046	0.216	0.008	-0.001	-0.260
	(2.20)	(0.52)	(2.36)	(0.08)	(0.00)	(1.08)
Minority Language	-0.110	-0.206	-0.124	-0.205	-0.055	-0.271
Concentration	(8.77)	(15.25)	(9.42)	(14.53)	(0.81)	(3.90)
Linguistic Distance	1.690	-7.009	-1.102	-6.460	-4.655	-10.492
	(5.18)	(19.81)	(3.09)	(16.69)	(3.84)	(8.25)
Miles Origin Country	0.091	0.219	0.066	0.231	0.126	0.114
From Canada/1000[c]	(2.40)	(5.44)	(1.51)	(4.92)	(1.35)	(1.17)
Refugee	-0.944	-2.256	-0.938	-2.133	-1.143	-3.380
	(7.92)	(16.54)	(7.62)	(14.87)	(2.03)	(5.73)
Colony	0.948	0.754	0.960	0.592	1.552	2.402
	(7.54)	(5.57)	(7.01)	(3.97)	(3.18)	(4.81)
Sample Size	32,168		26,484		5,483	
c^2	13,432		11,801		1,861	
Pseudo R^2	0.2453		0.2598		0.2068	

(a) Excludes Atlantic Provinces: see text for explanation.
(b) Variable not relevant.
(c) Variable defined with reference to Montreal in the equations for Quebec.

Note: L1 = Speaks neither English nor French; L2 = Speaks an Official language but usually speaks a non-official language at home; L3 = Speaks an Official language and usually speaks an official language at home.
Asymptotic 't' statistics in parentheses.

Source: 1991 Census of Canada, Public Use Microdata File (Individuals)

Table 6

Predicted Distributions across Language Categories by Minority Language Concentration, by Duration of Residence[(a)]

Minority Lang. Concentration	Example[(b)]	After 5 years in Canada			After 15 years in Canada		
		L1	L2	L3	L1	L2	L3
0.0	--	6.25	71.89	21.86	1.65	56.25	42.09
1.0	Punjabi in Toronto	7.07	72.81	20.12	1.92	58.38	39.71
2.0	Arabic in Montreal	7.97	73.55	18.47	2.22	60.43	37.36
3.0	Portuguese in Toronto	8.97	74.11	16.92	2.55	62.39	35.06
4.0	Chinese in Sudbury	10.07	74.48	15.46	2.94	64.25	32.82
5.0	German in Kitchner	11.26	74.65	14.08	3.37	65.99	30.64
6.0	Italian in Windsor	12.57	74.63	12.80	3.85	67.61	28.54
7.0	Italian in Toronto	13.99	74.41	11.60	4.39	69.10	26.51
8.0	--	15.54	73.98	10.48	5.00	70.43	24.56
9.0	Chinese in Vancouver	17.20	73.36	9.45	5.68	71.62	22.70

Note: L1 = Speaks neither English nor French; L2 = Speaks an official language but usually speaks a non-official language at home; L3 = Speaks an official language and usually speaks an official language at home.

(a) = Predictions computed at means of all variables other than duration of residence. Row totals for each immigration period may not add to 100.0 due to rounding.

(b) Examples of minority language concentration values that are close to designated value. There is no close value for 8.0. Mean value is 2.2.

Source: Estimates in Table 5.

Table 7

Predicted Distributions across Language Categories by Miles Between Origin and Canada, by Duration of Residence[a]

Miles	Example of Origin[b]	After 5 years in Canada			After 15 years in Canada		
		L1	L2	L3	L1	L2	L3
2,000	Guadeloupe, WI	11.13	76.15	12.72	3.38	68.47	28.14
3,000	Bogota	10.09	75.56	14.35	2.98	66.11	30.90
4,000	Warsaw	9.11	74.74	16.14	2.62	63.58	33.80
5,000	Seoul	8.20	73.69	18.10	2.29	60.89	36.82
6,000	Jerusalem	7.36	72.41	20.23	2.00	58.07	39.94
7,000	Ho Chi Min	6.58	70.89	22.53	1.73	55.14	43.13
8,000	Bombay	5.86	69.14	24.99	1.50	52.13	46.38

(a) For notes to Table, see Table 6.

(b) Mileage approximate (within 10 percent) from designated city to nearest Gateway City (Vancouver or Toronto) in Canada.

Source: Estimates in Table 5.

Table 8

Predicted Distributions across Language Categories by Linguistic Score, Refugee Status and if Birthplace is a Former Colony, by Duration of Residence[a]

Variable		After 5 years in Canada			After 15 years in Canada		
		L1	L2	L3	L1	L2	L3
Linguistic Score	Language[b]						
1.38	Chinese	13.08	80.79	6.13	4.41	80.55	15.04
2.00	Polish	8.14	73.47	18.40	2.26	60.46	37.27
2.75	Dutch	5.48	62.30	32.22	1.29	43.42	55.28
Refugee Origin Status							
Refugee[c]		20.37	74.40	5.24	7.31	79.00	13.68
Other		8.22	73.71	18.07	2.30	60.93	36.77
Colonial Origin Status							
Colony[d]		4.33	78.93	16.74	1.21	64.91	33.88
Other		8.22	73.71	18.07	2.30	60.93	36.77

(a) For notes to Table, see Table 8.
(b) Language scores where Chinese is the lowest and Dutch is the highest for the languages identified in the data.
(c) Unity if from Vietnam or the former USSR.
(d) Unity if from South Asia, Vietnam, Africa, Hong Kong or the Philippines.

Source: Estimates in Table 5.