NBER WORKING PAPER SERIES

CULTURE AND LANGUAGE

Edward P. Lazear

Working Paper 5249

CULTURE AND LANGUAGE

ABSTRACT

Common culture and common language facilitate trade between individuals. Minorities have incentives to become assimilated and to learn the majority language so that they have a larger pool of potential trading partners. The value of assimilation is larger to an individual from a small minority than to one from a large minority group. When a society has a very large majority of individuals from one culture, individuals from minority groups will be assimilated more quickly. Assimilation is less likely when an immigrant's native culture and language is broadly represented in his new country. Also, when governments protect minority interests directly, incentives to be assimilated into the majority culture are reduced. Both factors may explain the recent rise in multiculturalism.

Individuals do not properly internalize the social value of assimilation. They ignore the benefits that others receive when they learn the majority language and become assimilated. In a pluralistic society, a government policy that encourages diverse cultural immigration over concentrated immigration is likely to increase the welfare of the native population. In the absence of strong offsetting effects, policies which encourage multiculturalism reduce the amount of trade and have adverse welfare consequences. Conversely, policies that subsidize assimilation and the acquisition of majority language skills can be socially beneficial.

The theory is tested and confirmed by examining U.S. Census data, which reveals that the likelihood that an immigrant will learn English is inversely related to the proportion of the local population that speaks his or her native language.

Edward P. Lazear
Graduate School of Business
Stanford University
Stanford, CA 94305-5015
and NBER

Culture is defined by anthropologists in a variety of ways. The definition usually includes some notion of shared values, beliefs, expectations, customs, jargon, and rituals. Language is the set of common sounds and symbols by which individuals communicate.[1] Societies may include a number of cultures and languages. For example, while the majority of Americans speak English, a significant proportion are Spanish speaking. Even among the English speakers, accents and vocabularies differ by region of the country.

Multiculturalism, or the tolerance by a society of many different cultures and languages, seems to be on the rise in the United States. This shows up in a number of ways. One of the most tangible of these is the recent growth of bilingual education. The idea that a child would be taught in his or her native language in public schools was alien to American ideals until recently. In fact, the situation was the reverse. Most immigrants insisted that their children be taught in English so that they could become "Americans." The growth of multiculturalism, for good or bad, takes the view that Americans speak many languages and have many different cultures.[2]

What accounts for these changes over time? When do immigrants hold on to their native cultures and languages? Under which circumstances is assimilation most likely to occur? Does assimilation occur at socially desirable rates? How do government transfer

---

[1]Both terms are somewhat ambiguous. "Culture" is the more amorphous of the two terms, but even "language" lends itself to somewhat blurred distinctions. Within any given language, there are dialects that belong to certain subgroups. Some of the words are sufficiently idiosyncratic that only individuals within that subgroup understand their meanings. Shelling (1960) was perhaps the first to recognize the importance of focal points. Culture and language are among the most general kinds of focal points.

[2]Sowell (1974) is a much more ambitious attempt to understand culture. He attempts to define the relation of culture to most aspects of social, economic, and political life on a global scale. This paper shies away from being specific on culture and how it operates.

policies affect assimilation? Is the localization of minorities into neighborhoods a natural outgrowth of maximizing behavior or are ghettos sustained only when the majority exerts pressure on the minority? Is chauvinistic behavior by some societies socially beneficial or merely an emotional response without any social value? These are a few of the questions that are addressed in what follows.

A Model of Culture

Trade between individuals is facilitated when all traders share a common culture and language. A common culture allows individuals to trade with one another without intermediaries. In the case of language, this is most clear. If two agents speak the same language, they can negotiate a contract without the use of a translator.[3]

The model presented here focuses on common culture and the facilitation of trade.[4] Let us begin by assuming that an individual randomly encounters one and only one other individual in each period.[5] Suppose that the expected value to one party of meeting another individual with whom one can trade is θ. Initially, suppose that there are only two cultures in a country, labeled A and B. Define p_a as the proportion of individuals who belong to culture A and p_b as the proportion of individuals who belong to culture B. The majority

---

[3]Adelstein (1994) points out that to use a common language, individuals must give up some freedom. They cannot retain the right to attach whatever verbal labels they like. To communicate, individuals must tailor their expressions to meet the expectations of others. He argues that language is not simply communication, but a way to coordinate behavior. These ideas fit well with the use of the term "language" throughout this paper.

[4]The empirical literature on the economic returns to assimilation began with Chiswick (1978). More to the point of this analysis is the work by McManus, Gould, and Welch (1983), which shows that english speaking hispanic americans do better in the labor market than non-english speaking ones. Also, Chiswick (1991) finds that both speaking and reading fluency affect earnings, with reading fluency playing the more important role.

[5]Lazear (1994) uses a random encounter model to analyze the choice by businesses of a corporate culture.

culture is A, which means that $p_a > p_b$. It is possible that $p_a + p_b > 1$ since one individual can belong to two cultures as, for example, in the case of bilingual persons. In order for trade to occur, an individual must encounter another individual with his own culture. If the value of a trade is θ, then the expected gains from trade that accrues to As and Bs are

(1) a. $R_a = \theta p_a$

and

  b. $R_b = \theta p_b$.

Since $p_a > p_b$, $R_a > R_b$. Individuals from the majority are richer than those from the minority.

Either type of individual can acquire the culture of the other group. When Bs acquire As culture, they become "assimilated" into the A group. They may still retain some or all of their old culture, but they now have the ability to trade with the majority group. In the case of language, this can be thought of as becoming fluent in the majority language, while retaining the ability to speak in the native tongue. Similarly, As might acquire the B culture or learn B's language.

It takes time to acquire the new culture or to learn the new language. Time has a cost, which is modeled as reducing the probability that a trading encounter takes place. Since a trading encounter has expected value θp_a to As and θp_b to Bs, the cost of acquiring the culture is

$$\theta p_a t_j$$

for As and

$$\theta p_b t_j$$

for Bs, where t_j is an individual specific time parameter that measures the efficiency with which individual j acquires the new culture. The density and distribution functions for t_j are denoted $g(t_j)$ and $G(t_j)$, respectively.

3

It is now straightforward to determine whether an individual will become bicultural. The gains to minority individual j from becoming assimilated are

(2)   Gains to B from assimilation $= \theta - \theta p_b$

$$= \theta(1 - p_b)$$

since bicultural individuals trade with certainty and Bs trade with probability p_b.

He becomes bicultural if the gains exceed the costs, or iff

(3) $\theta(1-p_b) > \theta p_b t_j$.

The l.h.s. of (3) denotes the returns to learning culture or language A, which is that an individual can now trade with A types in addition to B types. The r.h.s. of (3) is the cost of learning A, as described above.

Expression (3) can be rewritten as

$(1-p_b)/p_b > t_j$

so the proportion of Bs who will become assimilated is simply

(4)   Proportion of Bs who learn $A = prob(\dfrac{(1-p_b)}{p_b} > t_j)$

$$= G(\dfrac{(1-p_b)}{p_b})$$

Proportion $G(\dfrac{(1-p_b)}{p_b})$ of the Bs are sufficiently efficient at acquiring the new culture to make it worthwhile.

Analogously, the gains to individual j of group A from learning culture B are

4

(5)   Gains to A from learning B = $\theta(1-p_a)$

An A becomes bicultural if

(6) $\theta(1-p_a) > \theta p_a t_j$

or if

$(1-p_a)/p_a > t_j$.

Thus, the proportion of As who learn culture B is then

(7)   Proportion of As who learn B = $prob(\dfrac{(1-p_a)}{p_a} > t_j)$

$$= G(\dfrac{(1-p_a)}{p_a}) \ .$$

(This assumes that the distribution of costs of learning B is the same as the distribution of the costs of learning A. We use this as a benchmark assumption so that behavioral differences are not a result of differences in distributional assumptions across groups.)

This analysis now leads to the following propositions:

Proposition 1: The proportion of a minority group that becomes assimilated into the majority culture is decreasing in the proportion of the population comprised by the minority group.

Proof: This follows trivially from (4) since $G(\dfrac{(1-p_b)}{p_b})$ is decreasing in p_b.

5

Proposition 2: The proportion of the majority group that learns the minority language is decreasing in the proportion of the population that speaks the majority language. Further, the proportion of majority members who learn the minority language is smaller than the proportion of minority members who learn the majority language.

Proof: From (7), $\partial G((1-p_a)/p_a) / \partial p_a < 0$. Also, since $p_a > p_b$,

$(1-p_b)/p_b > (1-p_a)/p_a$, the proportion of Bs who learn A is larger than the proportion of As who learn B from (4) and (7).

Discussion:

As p_b decreases, the minority group becomes smaller relative to the majority, which means that random contact with others is more likely to result in an encounter with someone of a different type. Thus, when the majority is very large relative to the minority, minority individuals are very unlikely to encounter others of their own culture and language. Minorities must be assimilated in order to survive in the society. Conversely, if the majority holds only a slight population advantage over the minority, neither group has as much incentive to learn the other language and the likelihood that the majority learns the minority language is almost as great as the likelihood that of the minority learns the majority language.

A member of the majority might learn the minority language. Individuals who are particularly adept at learning languages or cultural traditions quickly are have low values of t_j. Consequently, some majority members will learn the minority language even though encounters with minority members are relatively rare.

Note that the relevant group for analysis is formally defined as that set of individuals with whom random encounters are equally (and most) likely. An individual's circle of

encounters depends on the amount of specialization in the economy. In a society with no specialization where each individual produces everything that he consumes, there is little reason to encounter individuals outside a very narrow group of family and neighbors. Indeed, as countries have increased their levels of specialization and trade with other countries for basic goods, they have found it valuable to learn a common language, which in the current world is English. Air traffic control provides another good example. Were all flights domestic ones, there would be no need to communicate in a common language. But with the increase in international travel it becomes necessary to use one language for air traffic control - again the choice was English.

Externalities:

When an individual decides to learn the language of the majority, he takes into account only the value to himself of the enriched trading opportunities. But there is positive externality generated each time an individual learns another culture or language. When a Spanish speaker in the U.S. learns English, not only does he open up opportunities for himself, but he also makes it possible for English speakers to trade with him, which confers a benefit on the majority.

As a result of this externality, the rate of assimilation is below the socially optimal level. Too much multiculturalism persists when individuals take only their own interests into account. Societies may push, therefore, to prevent the spread of multiculturalism. Some states have introduced legislation to make English the official language. This is an attempt to use the government to raise the costs of multiculturalism, presumably with the purpose of internalizing the externality. The formal analysis follows.

The gains to a B of learning A is given by (2). When a B does not know language A,

a randomly encountered partner receives an expected value of

$$\theta p_b$$

Proportion p_b of the population speak B and can trade with this individual. Trade yields θ to each of the trading parties.

If B learns A, trade occurs with any randomly encountered partner so the randomly encountered partner receives θ. The difference is

$$\theta(1-p_b) \, ,$$

which is the size of the externality. The individual sees his returns as $\theta(1-p_b)$. The total social returns are actually

$$2\theta(1 - p_b) \, .$$

Thus, the individual takes into account only half of the returns, i.e., he ignores his trading partner's gain.[6]

How much of the externality is internalized? To the extent that most trades proceed through markets, it might be expected that prices would adjust to eliminate the externality. For example, it is known that immigrants who are fluent in English earn higher wages than

[6]This is analogous to the inability to take proper care that is part of the law and economics literature. See, for example, Green (1976) whose point is very close to the one made here. A similar point arises in the literature on networks and standards. Recently, Jackson and Wolinsky (1994) have analyzed whether efficient networks are stable and find that they are not under a variety of circumstances. Finally, Laitin (1994) treats the use of language as a coordination game, much like the standards literature in economics. See also Pool (1991).

There is no symmetric externality reducing costs. When an individual takes time t_j out of the market to learn the new language, he reduces the number of individuals in the minority relative to those in the majority. That is, p_b declines and p_a rises. While the effect is small, this is actually socially beneficial because the probability that a match occurs increases as the number of minority members declines. This merely reinforces the externality. This result depends on the match technology. If the number of encounters goes up with the number of individuals in the population, then the argument is more complicated and under certain matching technologies, might work to offset the externality on the benefit side. This relates to issues of economies of scale. If losing a trader in no way affects the utility of other traders, then scale economies have been exhausted. This is discussed below in the context of ghettos.

those who are not.[7] Does this eliminate externalities? The answer is no for two reasons. First, some transactions do not operate through the market. Asking for directions from strangers is not a market transaction, but common language is essential. Still, these non-market transactions may not involve major social costs.

Second, even if all were internalized by one individual, there remains an inefficiency associated with double-sided moral hazard. To see this, note that the total social return to each match is 2θ. Suppose that a minority worker who learned the majority language captured the full $2\theta(1-p_k)$ of return in the form of higher wages. He would have appropriate incentives to learn the language. But then the employer would have insufficient incentives to learn the minority language, even if it were cheaper for the employer to learn. To the extent that matches could be set up in advance, this would not be a major problem. Fast learning employers would learn the minority language and would hire only B speaking minority members. Fast learning minority members would learn the majority language and would work for strictly A speaking employers. The stochastic nature of matches prevents this from occurring with full efficiency. Indeed, the value of having national borders drawn so that everyone within them shares the same language is that it eliminates random encounters between individuals from different cultures. The cost is that some positive value from diversity and some economies of scale may be sacrificed.

Multiculturalism:

Suppose that instead of only two cultures, there are many distinct cultures.[8] How

---

[7]See, for example, the discussion of McManus, Gould and Welch (1983).

[8]"Distinct" culture is defined to mean that initially, no one belongs to more than one culture. A society where cultural change is occurring cannot have distinct cultures since some individuals are bilingual. This point is ignored in subsequent analysis in order to save on notation. Specifically, the

9

does the analysis change? One result, a generalization of proposition 1, follows immediately:

Proposition 3: Let there be i+1 cultures, where a is defined to be the plurality culture. Then the probability that an individual from culture i will be assimilated into culture a is decreasing in the proportion of individuals in culture i.

Proof: Denote the proportion of individuals who are of the majority type as p_a. A generalization of (3) reveals that an individual from group i will learn A if $\frac{p_a}{p_i} > t_j$, where p_i is the proportion of individuals who know culture i. As before, the proportion of i speakers who learn A is given by $G(\frac{p_a}{p_i})$, which is decreasing in p_i.

Proposition 3 implies that balanced immigration leads to a situation where all individuals are more likely to learn the majority language in order to communicate with one another. In countries like India, where there are a very large number of languages spoken, minority cultures learn English (or Hindi) so that they can communicate with the vast majority of individuals in their society who do not speak their language.[9] Balanced immigration should be defined in terms of balance in the receiving country. At issue is that each minority is a small part of the total in the new country. This means that individuals

---

assumption permits us to be unconcerned about the proportion of individuals who are bilingual or multilingual in each of the i languages.

[9]Chiswick and Miller (1994b) find that in Canada, immigrants tend to learn the language that predominates in their region of residence.

10

from populous countries have lower probabilities of being admitted to the new country than those from smaller countries. Unless the receiving country wanted, for some reason, to be a microcosm of the entire world, immigration quotas would be stated in terms of equal absolute numbers, not in terms of proportions of the donor country populations.

The idea generalizes. The more equal are the proportions in a society, the less likely is any one individual to switch to another language. Countries where there are a number of groups of similar size may find this configuration stable for a long period of time. Belgium is a case in point. Belgium, which has significant proportions of both Flemish and French speakers, has remained a dual language country. In fact, the country is becoming more separate as the rules of localities and ethnic groups take precedence over federal rules.

Although the proportion of individuals who learn another language may decline when the number of different minority languages increases, another equilibrium is possible. The society may end up gravitating to one common language, which is used by all. Israel made a conscious decision to use Hebrew as its language and to set up programs to teach the vast majority of immigrants this language. Since the immigrants came from many different countries and spoke different languages, (e.g., German, Yiddish, Russian, Romanian, Hungarian, Polish, French, Arabic, and English), p_i was small relative to the Hebrew speaking proportion. As a result, Israel was successful in inducing its population to use a language that was not spoken by any group of people for 2000 years.[10] Indeed, many now fear that the huge wave of Russian immigrants who have entered Israel in recent years may have reached a critical mass. Rather than becoming assimilated into Israeli culture, the new

---

[10]The former Soviet Union provides another example. Because of the many different languages spoken throughout the USSR, russian was made the official language and the one that Czechs, Poles, and others had to learn in school.

Russian Israelis may continue to speak Russian and subscribe to Russian culture.[11]

This point has been made before in another context. McNeill (1976) writes,[12]

> From the 1830s and especially after 1850, rapid urban growth together with the ravages of a new disease, cholera, disrupted cultural patterns of long standing in the Hapsburg monarchy. Peasant migrants into the towns of Bohemia and Hungary had long been accustomed to learn German, and in a few generations, their descendants became German in sentiment as well as in language. This process began to falter in the nineteenth century. When the number of Slav- and Magyar-speaking migrants living in the cities of the monarchy passed a certain point, newcomers no longer had to learn German for everyday life. Presently nationalist ideals took root and made a German identity seem unpatriotic. The result was that Prague became a Czech- and Budapest a Magyar-speaking city within half a century.

Multiculturalism imposes a direct cost on a society. The expected per person surplus in a purely multicultural society (where no one is bicultural) is

$$\theta \sum_i p_i^2$$

since type i encounters another type i p_i^2 of the time. Each encounter yields surplus θ to each side. Surplus is maximized when $p_a = 1$. Unicultural societies allow the most trade to occur. All encounters are matches. In a multicultural society, individuals suffer when they cannot deal with differently cultured individuals.

This result may be offset if different cultures bring enriched trading opportunities that would be absent in a single culture society. To take a somewhat trivial example, the United States has a wide variety of cuisines because of its underlying cultural diversity. Being able to choose among many different kinds of restaurants provides more social value than would be present were only English food available. Formally, this amounts to saying that θ is higher in multicultural societies than it is in single culture societies. If this effect is

---

[11]Chiswick (1993) studies the acquisition of Hebrew language skills in Israel. As in the US, Chiswick finds that the ability to speak the majority language increases earnings in Israel.

[12]McNeill (1976), p. 57.

sufficiently large, it can swamp the reduced surplus that comes about from the failure to match types as frequently.[13]

Open Immigration:

Before an individual decides to move to another country, he takes into account that the culture that he will encounter is different from the one with which he is familiar. This affects the decision to migrate. The more foreign the new culture, the less likely an individual is to migrate. Does this force correct the externality problem mentioned above? In a nutshell, the answer is no.

Consider an immigrant of type i who is considering a move to a new land in which everyone is unicultural. When she arrives, she can either learn the plurality language, say, English, or she can simply live as an i in a primarily English speaking country. If she moves, but remains an i, her expected rent is $p_i\theta$. If she moves and assimilates, her expected rent is $\theta(p_a+p_i)$ - $\theta p_i t_j$. By bearing cost $\theta p_i t_j$, individual j can become bicultural and increase her number of trading partners. Alternatively, she can migrate and simply remain unicultural. She moves to the new country if

(8) $\text{Max} \{ \theta p_i , [\theta(p_a+p_i) - \theta p_i t_j] \}$ - moving costs > R

where R is the rent that she receives by trading in her home country.

Define t* as the solution to

---

[13]The value of diversity has been modeled formally by Weitzman (1992). The desire for diversity is the opposite of the desire for commonality that motivates this analysis. But some, see DiPasquale and Glaeser (1994), have argued that there are direct costs of having a multicultural society. They analyze riots and find that the likelihood depends on ethnic heterogeneity.

13

$$p_i\theta = p_a\theta - \theta t^*$$

or

$$t^* = \frac{p_a}{p_i} \, .$$

Every immigrant whose $t_j < t^*$ chooses the second branch in (8) and learns the new language.

A generalization of the earlier discussion shows that every immigrant ignores θp_a of the returns to learning the majority language. Thus, as long as there are any individuals who have $t_j > t^*$, but also meet the condition of (8), there will be an externality. Expression (8) merely requires that the sum of moving costs and rent in the former country fall below the level that can be obtained in the new country under the best of either switching or staying with the old language. Any immigrant who satisfies (8) and also has $t_j > t^*$ fails to learn the plurality language of the new country, but comes anyway.

Individuals are more likely to move from countries that have large proportions of compatriots in the new land and when they arrive, they are less likely to switch languages than those from other countries. For example, there is a significant number of Spanish speakers, but many fewer Hungarian speakers in the United States. The rents that accrue to a Spanish speaker in the United States are θp_s, and to a Hungarian speaker are θp_h, where the s and h denote Spanish and Hungarian, respectively. Since $p_s > p_h$, a Spanish speaker's rents are higher. This implies that Spanish speaking individuals with sufficiently high costs of learning English to prevent switching are more likely to migrate to the United States than a similarly situated Hungarian. Furthermore, proposition 2 implies that once here, the probability that an immigrant will switch to English declines as the proportion of her compatriots in the new country's population rises.

14

It is now possible to evaluate the welfare implications of admitting immigrants from various cultures. First note that the value of an immigrant to the native population is given by

(probability that the immigrant does not learn the plurality language)(value of matches with own group)

+ (probability that the immigrant does learn the plurality language)(net value of matches with plurality and own group)

or[14]

$$(9) \quad Value = [1-G(\frac{p_a}{p_i})]\theta p_i + G(\frac{p_a}{p_i})\theta(p_a+p_i)$$

$$= \theta p_i + \theta G(\frac{p_a}{p_i})p_a \quad for \ i \neq a$$

$$= \theta p_a \quad for \ i=a$$

The following propositions about the choice of immigrant native culture follow directly from (9)

Proposition 4: Choosing an immigrant from the plurality culture maximizes social welfare when the plurality is large relative to minority cultures and when minorities costs of learning the plurality language are high.

Proof: Utility of the native population is maximized by choosing an immigrant from the plurality culture when

$$\theta p_a > \theta p_i + G(\frac{p_a}{p_i}) \theta p_a$$

---

[14]Note that this calculation ignores costs of learning the new language, c_i, borne by the immigrant. As such, it reflects social value where society is defined to exclude the utility of any potential immigrant. An ex ante, rather than ex post, definition of native population is used.

15

from (9), or when

$$(1 - G(\frac{p_a}{p_i}))\, p_a > p_i \;.$$

This is most likely to happen when p_a is large relative to p_i and when $(1-G)$ is large. When the costs of learning new language are high, then $(1-G(x))$ is large for any given x. | | |

Proposition 5: **When the choice is between immigrants from two minorities, the welfare of the native population is maximized by choosing the immigrant from the smaller minority when and only when immigrants from the smaller minority are sufficiently more likely to learn the plurality language.** | | |

Proof: Consider the choice between two minorities with proportions p_1 and p_2 in the native population. Let G_1 be the proportion of immigrants from group 1 who learn the new language, i.e., $G_1 = G_1(\frac{p_a}{p_1})$. Similarly, $G_2 = G_2(\frac{p_a}{p_2})$. If the density of costs were the same across groups, then more 2s would learn the plurality language than 1s because $p_2 < p_1$.

From (9), the gain to admitting an individual of type 2 over an individual of type 1 is

(P5.1) $(G_2 - G_1)p_a + (p_2 - p_1)$

The second term is negative. In order to favor the smaller group, the first term must be sufficiently positive to offset the first term. This requires that G_2 exceed G_1 by a sufficient margin. | | |

16

Consider the example above, where there are fewer Hungarian speaking Americans than there are Spanish speaking Americans. The difference in proportions p_i means that the Hungarian immigrant is more likely to learn English than the Spanish speaking immigrant. This is reflected in the first term of (P5.1) which favors Hungarian immigration. But very few Americans can trade with a Hungarian immigrant who opts against learning English; there is a larger number of Americans who can trade with a Spanish speaking immigrant who does not learn English. This effect is reflected in the second term of (P5.1), which is positive and favors Spanish speaking immigration. In order to favor an Hungarian over a Spanish speaking immigrant, it is necessary that the Hungarian is enough more likely to learn English than the Spanish speaker to outweigh the advantage to the larger Spanish speaking community of having another Spanish speaker in the American population.

Balanced immigration is favored when immigrants from small minorities are sufficiently more likely to learn English than those from the larger minorities. Some numbers are useful. In the U.S., Spanish speakers comprise the largest minority, but only around 2% of the current U.S. population is exclusively Spanish speaking. Thus, an upper bound to $p_1 - p_2$ is .02. Estimates provided below reveal that $(G_2 - G_1)p_a$ is quite large[15], swamping the .02 figure. Thus, balanced immigration is beneficial from the viewpoint of the native U.S. population.

When immigration is slow and balanced so that there are very few others in the society who share one's culture, assimilation is relatively rapid. When large groups immigrate at one time, many other individuals in the new country share the immigrant's native culture and language, which reduces the gains to adopting the new country's majority

---

[15]Fluency among Spanish speakers is .55 and .80 among non-Spanish speaking immigrants. The proportion of the native population who are fluent in english is around .95. One estimate of $(G_2-G_1)p_a$ is therefore .95(.25) = .24, which exceeds .02 by an order of magnitude.

17

culture.[16]

Irrespective of whether large minorities or small minorities are admitted, the welfare maximizing strategy is to admit immigrants who will be assimilated quickly. Then everyone in society gets the advantage of being able to trade with the new immigrant. When immigration is sufficiently slow and/or balanced, an immigrant who does not learn the new country's language cannot trade with anyone.

Selling Immigration Slots:

Is there a price mechanism that induces the value maximizing immigration? The following proposition provides the answer:

Proposition 6: An entry fee of F to all immigrants generally induces inefficient immigration from the native population's point of view. In particular, too many immigrants from large minority groups who will not learn the plurality language immigrate.

Proof: Propositions 4 and 5 imply that the native population prefers immigrants who learn the plurality language to immigrants who do not learn the plurality language. As before, denote the large minority proportion p_1 and the small minority proportion p_2. Denote the rent from remaining in the old country R_1 and R_2, respectively. From (8), a member of the large minority migrates and retains his old language when

(P6.1) $p_1\theta > R_1 +$ moving costs $+ F$

and when

---

[16]Immigrants do not necessarily prefer balanced immigration. Having a large group of individuals from one's own background in the recipient country reduces moving costs and increases trading opportunities.

(P6.2) $\quad t_j > \dfrac{p_a}{p_1}$

hold.

A member from a small minority migrates and learns the new language when

(P6.3) $(p_a + p_2)\theta - p_2\theta t_j > R_2 +$ moving costs $+ F$

and when

(P6.2) $\quad t_j < \dfrac{p_a}{p_2}$

hold. One cannot guarantee that an F exists such that large minorities who retain their language will not immigrate, but small minorities who learn the plurality language will immigrate. For p_1 close to p_a or for t_j sufficiently large, there is no F such that (P6.3) and (P6.4) hold when (P6.1) and (P6.2) do not. Thus, members from large minorities immigrate but do not switch languages and members of small minorities do not immigrate at all. Under varying parameters, this reverses the preference ordering of the native population. | | |

The problem is that immigrants are willing to pay a high price for entry when their situation at home is very bad and when they can do well even without learning the majority language. Neither of these reasons aligns with the preferences of the native population of the receiving country. In particular, the native population always prefers that an immigrant, irrespective of origin, learn the plurality language.

Because an entry fee is not sufficient to induce desired immigration, other conditions must be attached to immigration. One way to remedy the problem pointed out in proposition

6 is to make the entry fee contingent on learning the plurality language. A refund could be given to those who pass a language test within a period of time. This kind of rebate structure is likely to be opposed very strongly by native members of large minority groups, whose compatriots are the individuals who were most likely to immigrate and retain the old language.[17] Alternatively, immigrants could be required to pass a language test before citizenship would be awarded.

Borjas (1985, 1994) has argued that the recent waves of hispanic immigrants have been of lower quality than earlier waves. This result is a direct implication of the model. As the proportion of Spanish speaking individuals in the United States rises, the Spanish speakers who previously were unwilling to come to the U.S. because their language switching costs were too high now find it worthwhile to come and simply not switch languages. To the extent that ability to learn English is positively correlated with other learning skills and/or ability, the Borjas finding of lower quality immigrants over time would be predicted.[18]

It is unnecessary to appeal to differences over time in the ability distribution of migrants to obtain this result. Even if the underlying ability distribution of immigrants were time invariant, incentives are different for newer cohorts of immigrants. A large population that speaks the home country's language reduces the incentives to switch so an immigrant of a given ability is less likely to switch if he is in a later cohort of immigrants. Thus, earnings growth and assimilation are likely to be slower for later cohorts, ability issues aside.

---

[17]The size of the entry fee is depends on the optimal flow of immigrants, which depends in turn on the importance of scale economies, crowding, agglomeration and other factors that are ignored in this model.

[18]This is not an automatic result of immigration, however. It depends crucially on the rate of immigration. As immigration rates slow, it is possible that the spanish speaking proportion of the U.S. population would decline, leading to less incentive for high switching cost individuals to migrate.

Extensions

Regionalization and Ghettos:

Concentration of minority populations into localities reduces minority incentives to become assimilated into the majority culture. Formally, p is defined in terms of the others with whom an individual comes into contact. If an individual travels only within local circles, the relevant population proportions may be very different from those of the country as a whole. Thus, while Polish may be spoken by a very small minority of Americans, there is a very large and concentrated Polish community in Chicago. A native Pole who immigrates to the Polish neighborhood in Chicago can function much better without a knowledge of English than he could were he to move, say, to San Francisco.

Ghettos are an extreme form of concentration. Societies sometimes force their minority members to live in ghettos and prevent them from mingling with the rest of society. Why might societies engage in this behavior? After all, minority members should choose to be in localities where their own group is over-represented as a voluntary action. But the construction of a ghetto sometimes involves coercion.

Such behavior can be rationalized in the current context. In the model as presented, a minority member would never choose to live outside a ghetto. Ghettos increase the probability of trade for minority members. Something else is needed to provide a motivation for minorities to leave the ghetto. Suppose that as a result of economies of scale, specialization, or access to wider international markets, trades that take place outside the ghetto are worth potentially more than those inside. Let trades outside be worth θ, as before, and let those inside be worth θ', with $\theta' \leq \theta$. For simplicity, consider the special case where there are only two distinct cultures, A and B. If communities are segregated, then every encounter by minority members and by majority members results in a match. The expected

per person social value is then

$$(10) \quad V_G \equiv p_a \theta + (1-p_a) \theta'$$

because p_a belong to the majority and $(1-p_a)$ to the minority.

If ghettos are destroyed and minority and majority individuals encounter each other randomly, then the expected per person social value is

$$(11) \quad V_{NG} \equiv p_a^2 \theta + (1-p_a)^2 \theta .$$

It can be shown[19] that $V_G > V_{NG}$ for all values of $\theta' < \theta$. Thus, the social value of trades in the ghetto economy is higher than that in the integrated economy. The reason is that moving individuals out of the ghetto improves the lot of the minority by less than it hurts the welfare of the majority who now encounter individuals with whom they cannot trade. Thus, majorities may try to segregate individuals into ghettos.[20]

---

[19]Proof: Δ, defined as $V_G - V_{NG}$, is increasing in θ'. Therefore, if Δ it is positive for all $p_a > 1/2$ when $\theta' = 0$, it is positive for any θ'. For Δ to be positive, it is necessary and sufficient for

$$p_a - p_a^2 - (1-p_a)^2 > 0.$$

The l.h.s. equals zero for $p_a = (1/2, 1)$. Taking the derivative of the l.h.s, with respect to p, one obtains

$$1 - 2p_a + 2(1-p_a) ,$$

which is positive at $p_a = 1/2$ and negative at $p = 1$. Thus, $\Delta > 0$ for $1/2 < p < 1$ so $V_G > V_{NG}$ for all values of θ, θ' and $p > 1/2$.

[20]Diamond and Maskin (1979) have explored the implications of different assumptions about random meeting probabilities for externalities. Specifically, I have assumed that movement of an individual from inside the ghetto to the outside does not increase the probability of an encounter (which is assumed to be 1). If changing the number of individuals outside the ghetto were to increase the likelihood of an encounter, the effect of the externality could be diminished.

It is possible that segregation may occur voluntarily, but there are also cases where minorities will not choose to be located in ghettos. Minority members prefer integration when $(1-p)\theta > \theta'$, which occurs when being segregated into ghettos significantly reduces the value of trade for minority members. But the solution that is always preferred by the majority is to keep minorities in ghettos. The fact that those who live inside ghettos may dislike segregation more than those who live outside them creates an obvious source of source of social tension.

There are important distributional consequences of ghetto policies. When minorities are forced into ghettos, they bear costs that exceed their benefits. This is clear since minorities are prevented from leaving the ghetto at rates that would be privately optimal, but socially inefficient. A direct consequence is that the majority enjoys a disproportionate share of the benefits from keeping minorities in ghettos.

Ghettos, which create distributional and other social problems, may be a result of unbalanced immigration. If immigration is slow and balanced across donor countries, then immigrants are more likely to acquire the culture of their new countries by proposition 1. As a result, ghettos are less likely to form.

The result that segregation into ghettos is efficient depends on two assumptions. The first is that economies of scale for the majority have been exhausted. Moving individuals from the ghetto to the outside does not increase the average value of a trade outside. Formally, this means that θ is independent of the number of individuals who are not in the ghetto.[21] This ignores the value of diversity, mentioned above. If there remain untapped gains to diversity from integration, then it is easy to reverse the result that ghettos are

---

[21]It is not necessary to assume that scale economies within the ghetto have been exhausted. It has already been shown that ghettos are efficient under the previous assumptions. Moving more into the ghetto would only raise V_G relative to V_{NG}.

efficient. Specifically, if trades in the integrated economy yield θ^* to the majority instead of the θ earned by the majority in the segregated economy, then sufficiently large values of θ^* relative to θ can always yield $V_{NG} > V_G$. While possible, a society like the United States in the 1990s is likely to have enough crossover to allow for most of the gains to diversity to be realized even if most of the minority population lives in ghettos.[22]

The second assumption is that children who are raised in the ghetto are not sufficiently less likely to be assimilated into the majority culture than those raised outside. If this assumption does not hold, then an argument for integration can be made. This issue is addressed in the next section.

Intergenerational Aspects:

Immigrants who came to the United States in the early part of the twentieth century may not have learned English themselves, but were overwhelmingly committed to having their children speak English as their primary language. While this may have made family communication more difficult, altruism by a parent for his or her offspring may induce parents to take actions that would benefit their children.[23]

To see this, consider a two period version of (10) and (11). In period 1, there are adults and children. In period 2, the children become adults. Only adults trade.

---

[22]The issue of how many communities a society should be divided into is related to the question of how many separate markets should there be in an economy. This has been analyzed by Economides and Siow (1988), who argue that externalities may produce too many markets, much like too many languages and distinct communities.

[23]This is similar to the externality argument made in Lazear (1983). There, the issue was whether parents would acquire enough schooling or move to the appropriate community, given the interests of their children.

Suppose that a minority child raised in a ghetto has a probability λ of being assimilated into the majority culture, whereas a child raised outside the ghetto is certain to become assimilated. This means that $p_a + \lambda(1-p_a)$ of the population will speak the majority language in the second period. Let δ be the discount factor. Then, the two period version of (10) and (11) is

(10') $\quad V_G = p_a\, \theta + (1-p_a)\, \theta' + \delta\{[p_a + \lambda(1-p_a)]\theta + [(1-p_a)-\lambda(1-p_a)]\theta'\}$

and

(11') $\quad V_{NG} = (p_a^2 + (1-p_a)^2)\theta + \delta\theta$.

Now, whether V_G exceeds V_{NG} or not depends on the parameters. If δ equals zero so that no weight is placed on the next generation, then (10') becomes (10) and (11') becomes (11) and the previous proof that $V_G > V_{NG}$ holds. But at the other extreme, suppose that $\delta=1$ (no discounting of the future), $\lambda=0$ (children raised in the ghetto never become assimilated) and $\theta'=0$ (within ghetto trades have zero value). Then, (10') becomes

$\qquad V_G = 2\, p_a\, \theta$

and (11') becomes

$\qquad V_{NG} = (p_a^2 + (1-p_a)^2)\theta + \theta$.

It is easily shown that V_{NG} now exceeds V_G for all p_a.[24]

---

[24] It is true that
$$(1-p_a)^2 > 0$$
or
$$1 - 2 p_a + p_a^2 > 0$$
so
$$p_a^2 + 1 - p_a > p_a$$
which implies that
$$2\theta p_a^2 + 2\theta - 2\theta p_a > 2\theta p_a$$

or that
$$p_a^2 + (1-p_a)^2 + \theta > 2\theta p_a \, .$$
Thus,

25

Indeed, as λ falls, the value of integration rises. Since $\partial V_G/\partial\lambda > 0$, and since V_{NG} is independent of λ, decreasing λ implies that V_G falls relative to V_{NG}. If children who are raised in the ghetto are less likely to be assimilated into the majority culture, then integration is more valuable.

It is also true that social welfare unambiguously increases in λ. No matter how much of the economy is in ghettos, any costless increase in λ always improves social welfare. Since $\partial V_G/\partial\lambda > 0$, facilitating assimilation of minority children into the majority culture has social benefits. This is the argument against bilingual education. To the extent that teaching a child in his or her native language reduces the likelihood of assimilation, bilingual education decreases social welfare and does so by hurting those that it purports to help.

The counterargument is that bilingual education allows children to learn other skills more rapidly, thereby increasing θ'. But this effect is likely to be swamped by the gains to assimilation. Even if childhood fluency required one full year of intense immersion in English language education, at worst this would simply shift a minority child's human capital investment profile back one year. Spending an extra year in English language kindergarten would mean that the child would graduate from high school at 19 rather than at 18. But from then on, earnings would be the same. Consider the two possibilities. If the minority high school graduate becomes fluent in English, he earns W from age 19 to age 65. If he retains the minority language, he earns γW from age 18 to age 65, where $\gamma<1$. Assuming that the costs of English immersion kindergarten are no greater than the costs of providing bilingual education throughout the entire school career, then English immersion kindergarten undertaken at age 5 is socially worthwhile if

---

$V_{NG} > V_G$.

$$\gamma W \int_{13}^{60} e^{-rt} dt \quad < W \int_{14}^{60} e^{-rt} dt$$

If r, the real rate of interest, equals .01, the condition holds as long as $\gamma < .97$. For r=.05, the condition holds as long as $\gamma < .95$. If non-assimilated minority members earn less than 97% of what assimilated individuals earn, then at a real rate of 1% it pays to avoid bilingual education to instead invest in English immersion kindergarten. Estimates of γ are well below .97. McManus, Gould and Welch (1983) find γ in the range of .65 to .80, depending on the skill level of the individual in question. There is little doubt that the economic calculation argues strongly against bilingual education.

It is no secret that American blacks have as a group been unable to escape the "culture of poverty."[25] Blacks have also found themselves concentrated in urban ghettos, at least during the second half of this century. The model in general, and this section in particular, are consistent with this phenomenon. To the extent that blacks have been isolated, they are less likely to choose assimilation and their children are less likely to be able to do so.

Government Transfers:

In the same way that local neighborhoods can offer support to minority individuals who are not members of the majority culture, so too can a government reduce the incentives to become assimilated. When an individual can go to the government for support, the value of changing culture is reduced. Formally, a government transfer program places a floor on

---

[25]These terms were used early by sociologists who studied chronically disadvantaged groups. See, for example, Oscar Lewis (1966) and Elliot Liebow (1967).

the amount that an individual can receive. In the absence of government transfers, the condition for learning the majority language is that

$$\theta(p_a + p_i) - \theta p_i t_j > \theta p_i,$$

(which reduces to $t_j < \dfrac{p_a}{p_i}$). A government transfer program can be thought of as

guaranteeing some average level of surplus, S. If

$$S > \theta(p_a + p_i) - \theta p_i t_j > \theta p_i,$$

an individual who would have learned the majority language in the absence of the government transfer will opt against learning it and will merely accept the transfer. Reducing the size of government transfers would increase the rate of assimilation in the society.

Chauvinism:

Certain societies have often been labeled "chauvinistic" because of their aggressive policy to keep their culture intact and to exclude alternative cultures. In North America, the most vivid example is in Quebec, where laws were passed to limit the use of English. This may be completely rational and perhaps socially efficient behavior. Suppose that in Quebec, everyone speaks French initially. Now suppose that through immigration, the proportion of French speaking individuals declines. There is an immediate cost to unilingual French speakers, who now encounter other individuals with whom they cannot speak. But there are secondary effects. Although the English speaking minority has more incentive to learn French than the French speaking majority does to learn English, some quick learning French

speakers may learn English. Besides increasing the desire of English speaking Canadians to move to Quebec, learning English imposes a direct costs in the form of time spent learning the language. It is conceivable that the new equilibrium might have a largely bilingual society or even a completely English speaking one. In the process, each individual has had to bear cost $\theta p_i t_j$ to make the switch.[26] The new equilibrium may be no better than the old; communication in French may be as effective as communication in English. The change has come about only at large social cost.[27] Thus, the chauvinistic policy may be efficient.[28]

The tradeoff can be shown formally. Imagine that initially Quebec consists of N individuals, all of whom are French speaking. The per person social value is simply θ. Now let M English speaking individuals migrate in. Define p_a as $\dfrac{N}{M+N}$ and Z, the total population as M+N. If all minorities are forced to learn French, then the per person social value is

$$(12) \quad \theta - (1-p_a)Z \int_0^\infty t_i g(t_i) dt_i$$

The first term reflects the value of trades. Since all speak French, the per person trade value is θ. The second term is the amount of cost borne by the new immigrants who must learn

---

[26]This is related to Schelling's (1978) famous tipping result, where neighborhoods change because of externalities imposed on other residents.

[27]A similar point is made by Lang (1986).

[28]An inefficiency of chauvinist policy is that it reduces the ability of Quebecois to trade with English speaking Canadians.

the new language.

If immigrants are not forced to learn the new language, some will do so voluntarily. They are those for whom

$$\frac{p_a}{1-p_a} > t_j.$$

Similarly, a smaller proportion of the majority learn English. They are those for whom

$$\frac{1-p_a}{p_a} > t_j.$$

Thus,

$$MG\left(\frac{p_a}{1-p_a}\right)$$
–

Voluntarily learn French and

$$NG\left(\frac{1-p_a}{p_a}\right)$$

voluntarily learn English. It is quite possible that the number of English speakers who learn

French, $NG(\frac{1-p_a}{p_a})$, exceeds the number of French who learn English, $MG(\frac{p_a}{1-p_a})$, even

though $p_a > 1/2$, because $N > M$.

The per person social value when individuals are not forced to learn French is

$$(13) \quad \theta\left[\left(\frac{N+MG\left(\frac{p_a}{1-p_a}\right)}{Z}\right)^2 + \frac{\left[M+NG\left(\frac{1-p_a}{p_a}\right)\right]^2}{Z}\right]$$

$$-\left\{\theta p_a Z \int_0^{\frac{1-p_a}{p_a}} t g(t_i)dt_i + \theta(1-p_a)Z \int_0^{\frac{p_a}{1-p_a}} t g(t_i)dt_i\right\}$$

The first term reflects the total value of trades in the economy and the second term reflects

all costs borne by both French and English speakers who learn the other language.

Chauvinism is a good policy if the expression in (12) exceeds the expression in (13).

The first term of (12) clearly exceeds the first term of (13). Every match is a trade in the

chauvinistic society. The absolute value of the second term in (12) is smaller than the

absolute value of the second term in (13) when N is large relative to M and when the high

31

cost learners among the minority are not too high cost. By forcing everyone to learn, society loses by the fact that some English speakers who learn French only at very high cost are forced to learn. By allowing bilingualism, these high cost English speakers are replaced by low cost French speakers. But since there are so many more French speakers, the total cost may be lower with a policy of chauvinism.

This provides a rationale for a chauvinistic policy: It is sometimes better to have a few bear costs, even if they are very high, than to have many bear some costs, even if they are quite low. Doing so trades off costs by one group against costs by another and takes into account different benefits as well.

Encouraging bilingualism is the opposite of chauvinism. A society that encourages bilingualism is willing to lose the advantages of a common language in order to reduce the costs of a language switch by one sub-group of the population. Because of the externality discussed above, however, there is too much desire for bilingualism. Individuals who speak the minority language may wish to be taught in their native tongue even though doing so reduces the number of individuals with whom they can trade. They ignore the social costs, borne by others, of maintaining their native language.

To encourage minority individuals to learn the majority language, chauvinistic policies can be used, which implicitly tax minorities. Alternatively, learning the majority language can be subsidized to offset the externality. This policy has been used successfully by the Israelis, who subsidize the teaching of Hebrew to immigrants.

Pioneers and Entrepreneurship:

The argument made earlier, that the first immigrants from a particular country are

likely to be more able than those from a later generation, can be generalized to a theory of pioneering. Just as the new immigrants must be low learning cost individuals, so to are entrants into any new activity the relatively low cost individuals. The Americans who first moved West last century were the ones who were best able to blaze trails and deal with a hostile environment. To the extent that those skills are correlated with underlying ability, the theory implies that pioneers are of higher ability. The migrants who came later could take advantage of the infrastructure created by the pioneers and could be "higher t_j" individuals.

The same argument applies to individuals who move into a new industry. To enter a new industry, an individual must learn a new "language" because everyone in the new industry speaks the new language. As more ordinary individuals move into the new industry, one can communicate and therefore trade with others even without knowing all of the technical jargon. Again, to the extent that these learning skills are correlated with some underlying ability, the entrepreneurs tend to be the most able.

Also, as before, even without appealing to differences over time in the inherent ability of migrants, those who come into new industries later will be less likely to switch to the new industrial culture. Thus, learning and earnings growth is likely to be lower for more recent recruits than for the earlier recruits, even if their abilities are the same.

Superior Cultures:

It has been assumed that an equilibrium where 100% speak English is no better nor any worse than one where 100% speak Spanish. But there may be factors which cause one language be more efficient. For example, English has many irregular cases and is less consistent phonetically than Spanish. As a result, the distribution of costs in going from

Spanish to English may lie to the right of that in going from English to Spanish. This could cause a society to favor a Spanish-only policy, even if the majority speaks English.

It could also be the case that some languages and cultures are simply more efficient than others. If it were easier to communicate using an alphabet than with Chinese characters, then θ might be higher for trades conducted in a European language than for those in a Chinese or Japanese. Contracts might be easier to write and communication might be richer in one language than in another. Again, this could lead to favoring one language over another, even if the language were not spoken by the majority.

Persecuted Emigrants:

In some situations, individuals fled their native lands under extreme persecution. Russian Jews who fled the pogroms at the turn of the century, South Vietnamese who escaped in the early 70s before the communists took over, and the earliest settlers of the American colonies are some obvious examples of individuals who left their former lands under duress. Persecuted peoples may have been less attached to their former cultures than those who were not persecuted. It seems reasonable that persecuted individuals are more likely become assimilated than those who were not. Formally, this can be interpreted as the distribution of costs among the persecuted being stochastically dominated by other immigrants. Since $G(p_s/p_i)$ become assimilated, and since this is a larger number for the persecuted than for other immigrants, those who emigrate under duress are more likely to adopt the new culture.

Empirical Analysis

The theory is tested using data from the 1990 Census, 1/200 sample, which is a randomly selected half of the 1/100 sample. The sample for analysis consists of individuals who reported that they were born outside the U.S. Their probability of speaking English was studied.

The key variable is the proportion of the local population, measured by county, that is comprised of other individuals who speak a given immigrant's language. The measure used is

CNTYPCT = number of individuals in the county who were born in the same country as the given individual / number of individuals in the county

The theory says that larger is the minority proportion, the less likely is the individual to learn English.

The census contains a question on ability to speak English. That variable, English, is defined as:

English = 1 very well; 2 well; 3 not well; 4 not at all

A new variable, "FLUENT" is defined as corresponding to a value of 1 or 2 on the English question.

The sample was divided into those who were "fluent" and those who were not. The means of all variables are reported in table 1. Table 2 reports the means of CNTYPCT are by FLUENT and by ENGLISH.

As is obvious, the mean level of CNTYPCT are higher among the non-English speaking subsample. The correlation is clear. Individuals who are fluent in English are more likely to live in communities where smaller proportions of the other residents speak the

language of their native countries. Similarly, CNTYPCT is monotonically decreasing in the degree of fluency. The mean is lowest among those who speak English very well (English=1) and highest among those who do not speak English at all (English=4).

This evidence is useful, but not compelling. In particular, it may be that individuals who live in communities dominated by other immigrants from their own countries are more recent immigrants. Table 3 and 4 provide additional evidence.

Two additional variables are introduced in the logit of table 3. "IMMIG" refers to the date at which immigration occurred with higher values referring to earlier immigration. A value of 10 corresponds to arriving before 1950 and of 1 refers to arriving this year. AGE is the individual's age in 1990.

The earlier the individual migrated to the United States, the more likely he is to know English. Also, given year of migration, the older is the individual, the less likely is he to learn English. Both make sense. The older is an individual at the time of entry, the smaller the gain from learning English because the returns accrue over a shorter period.

Even when these factors are held constant, CNTYPCT still enter negatively. Furthermore, the coefficients are large and economically meaningful. Table 3 implies that a one standard deviation increase in CNTYPCT would decrease the probability of being fluent by about 12%, or about 1/5 the level of fluency for the overall sample.

Table 4 performs an ordered logit on the English variable. The results are similar. Again, raising CNTYPCT has a significant effect on knowing English. Here, a one standard deviation increase in CNTYPCT raises the probability of not knowing any English by about 5%. Only 11% of the sample knows no English at all, so this is a large effect.

The analyses of tables 2 through 4 cut across groups. But there are many individuals in the sample who speak either Spanish or Chinese. Specifically, 48% of immigrants speak

36

Spanish and 7% speak Chinese. The analysis was performed on each of these subgroups separately. The results are contained in tables 5 and 6.

Even within group, the effects remain strong. Among Spanish speakers, those who live in areas with many other Spanish speakers are less likely to speak English well. The same is true for the Chinese subsample.

It is interesting that the coefficients are so different between the Spanish speaking sub-sample and the Chinese speaking subsample. This may reflect a nonlinearity. Since the average value of CNTYPCT is higher for Spanish speakers than for Chinese speakers, this may merely reflect that the effect of CNTYPCT on fluency is concave. This was tested by entering $(CNTYPCT)^2$ in the logit. Although $(CNTYPCT)^2$ is positive and significant, the effects of CNTYPCT on fluency remain much larger in absolute terms for the Chinese speaking sub-sample.

Tables 1-6 provide compelling evidence that the effects predicted by the model are borne out in the data. Assimilation as proxied by English fluency is less likely in communities where there are a high proportion of individuals born in a given individual's native country.

Sorting v. Learning

While the effects are clearly strong, they are subject to another possible interpretation. It may be that individuals who are not fluent in English are more likely to move to areas where there are others of their own culture. This sorting explanation is clearly possible, but surely reflects the same mechanism described in this paper. Individuals who are not fluent in English move to areas that are dominated by their own culture precisely because they cannot trade with others unless they do. It is the gains from trade that influence this pattern

of location. If it were not necessary to be with other similarly cultured individuals to engage in trade, the locational pattern of immigrants would uncorrelated with English fluency.

Furthermore, even if those immigrants who know English are more likely to locate in areas with low proportions of compatriots, this does not negate the strategic learning explanation. The immigrant may know English when he arrived, but the decision to learn it and to migrate to the U.S. is likely to have been affected by his desire to trade with the larger community of Americans. Thus, the sorting story differs from the learning story primarily on the timing at which English was learned, not on the motives for learning English.

Finally, there is other evidence on these points. In a series of papers, Chiswick and Miller and Chiswick have documented similar results for a number of countries. Using an Australian data set, they find that the acquisition of destination language skills are inhibited by living in an area where many others speak the same minority language.[29] Chiswick finds that English fluency increases with the duration of time that the immigrant has been in the United States.[30]

Issues for the Future

There are many implications of the analysis that have not been tested. The most important relate changes over time. Earlier census data will shed light on whether rates of assimilation have changed over time.

Female immigrants, especially those out of the labor force, may have different encounter rates than male immigrants. If so, the value of learning the majority language

---

[29]See Chiswick and Miller (1994a, 1995).

[30]See Chiswick (1991).

may be lower to women and women would be less likely to learn the new language. It is possible to investigate this by stratifying the data by gender and by labor force status.

A similar point holds for occupation. Knowing the majority language is more important in some occupations than in others. The importance of communication is probably related to the skill level of the occupation. This can be investigated by examining other occupations.

Individuals may differ in their costs of learning a new culture or language on the basis of some observables. For example, more able individuals, proxied, say, by education level, may be more likely to learn a new language. This can be modeled and tested.

Finally, the number of individuals from a group, in addition to their proportion, may be an important determinant of the value of learning the majority language. Encounters may come more often in a dense neighborhood than in a sparsely populated one. As the number of encounters rises, the value of knowing the majority language rises as well. But as the number of encounters rises, the individual can take advantage of a smaller proportion of encounters and may find that he can satisfy his trade demand simply by dealing with individuals from his own minority. This can be investigated both theoretically and empirically.

Summary and Conclusion

Individuals from minority groups are more likely to adopt the culture and language of the majority when the minority group accounts for a small proportion of the total population. The incentives are greater for any individual to learn the majority language when only a few persons in the country speak his or her minority language. Thus, slow and balanced immigration, where the flow of individuals from any one culture is small, results

in more rapid assimilation than immigration that favors any one particular group.

Individuals from the majority may learn the language or culture of one of the minorities. But it is less likely that the majority will learn a minority language than that a minority will learn the majority language.

Empirical evidence from the 1990 U.S. Census demonstrates conclusively that immigrants are most likely to learn English when they live in communities that have small proportions of individuals from their own native country. Individuals who are from poorly represented groups learn English quickly. Those from groups with large proportions in the local population learn English more slowly. This is a rational response to the differences in the value of learning English across groups. The finding holds up within cultural groups as well as across groups.

Some additional points are summarized below:

1. Government transfers, which place a floor on consumption, reduce the incentives to adopt the majority culture and learn the majority language. The rise in government transfers, and the large group of current U.S. citizens and resident aliens who use Spanish as their primary language, may account for the increased taste for multiculturalism.

2. Individuals do not properly internalize the gains from assimilation. When an individual learns a new language, he trades off his personal benefits against his personal costs and ignores the benefits to his trading partner. As such, assimilation is too slow to maximize social welfare.

3. Immigrants to a given country are more likely to emigrate from countries whose cultures are well-represented in the population of the new country. Although multiculturalism reduces the overall value of trades within a society, once a country has a large minority, the value of introducing an immigrant from the large minority group may be

higher than that of introducing an immigrant from a small minority group. But if members of the smaller minority are sufficiently more likely to learn the plurality language, the native population favors admission of members of the smaller minority. This provides a case for balanced immigration.

4. There is no entry fee that guarantees efficient immigration. In particular, individuals who are from countries with large minorities in the new country are too likely to immigrate and maintain their own language. Immigrants from small minorities who would learn the new language were they to migrate are discouraged from doing so by the fee.

5. Bilingual education is almost certainly a poor investment.

6. Minority individuals who have the option of living in neighborhoods with many others who share their culture and speak their language are less likely to be assimilated. Elimination of ghettos is most valuable when individuals who are raised in the ghetto are much less likely than non-ghetto minorities to be assimilated into the majority culture. It is the effect of ghetto life on children that has the most negative social consequences.

7. Chauvinism, where a particular majority culture is tenaciously protected, may be rational and welfare increasing. The introduction of minority cultures can bring about a new equilibrium that is no better than the old. In the process of moving from the old equilibrium to the new one, transition costs are incurred.

References

Adelstein, Richard "Order and Planning," unpub. manuscript, Dept. of Economics, Wesleyan University, 1994.

Borjas, George "Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s?" NBER WP # 4866, 1994.

Borjas, George. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants", *J. of Labor Economics 3,* (October 1985): 463-89.

Chiswick, Barry. "The Effect of Americanization on Earnings of Foreign-Born Men," *J. Polit. Econ. 86,* (Oct 1978): 897-921.

Chiswick, Barry. "Hebrew Language Usage: Determinants and Effects on Earnings Among Immigrants in Israel." unpub. U. of Ill. Chicago, 1993..

Chiswick, Barry. "Speaking, Reading, and Earnings Among Low-Skilled Immigrants." *Journal of Labor Economics.* 9, 2 (April 1991): 149-170.

Chiswick, Barry and Paul Miller. "The Endogeneity Between Language and Earnings: International Analyses." Forthcoming, *Journal of Labor Economics.* 1995.

Chiswick, Barry and Paul Miller. "Ethnic Networks and Language Proficiency Among Immigrants." *Journal of Population Economics.* Forthcoming, 1994a.

Chiswick, Barry and Paul W. Miller. "Language choice among immigrants in a multi-lingual destination." *Journal of Population Economics.* 7 (1994b):119-31.

DiPasquale, Denise and Edward Glaeser " The L.A. Riot and the Economics of Urban Unrest," (Harvard University, mimeo, May 1994.)

Diamond, Peter and Eric Maskin "An equilibrium analysis of search and breach of contract, I: steady states." *Bell J. of Econ.* 10,1 (Spring, 1979): 282-316.

Economides, N. and A. Siow "The Division of Markets is Limited by the Extent of Liquidity (Spacial Competition with Externalities)" *Amer. Econ. Rev.* 78,1 (March, 1988), 108-21.

Green, Jerry R. "On the Optimal Structure of Liability Laws," *Bell Journal of Economics* 7 (Autumn 1976): 553-74.

Jackson, Matthew and Asher Wolinsky, "A Strategic Model of Social and Economic Networks," Northwestern U. mimeo, August, 1994.

Laitin, David D. "The Tower of Babel as a Coordination Game: Political Linguistics in Ghana," *American Political Science Review* 88 (3) (September 1994): 622-34.

Lang, Kevin "A Language Theory of Discrimination," *Quarterly Journal of Economics*, May 1986, pp. 363-82.

Lazear, Edward P. "Intergenerational Externalities," *Canadian Journal of Economics* 16 (May 1983): 212-28.

Lazear, Edward "Corporate Culture and the Diffusion of Values" (July 1994), forthcoming in conference volume from the Kiel Institute of World Economics.

Lewis, Oscar *La Vida; A Puerto Rican family in the culture of poverty - San Juan and New York*. New York: Random House, 1966.

Liebow, Elliot *Tally's Corner; a study of Negro streetcorner men*. Boston: Little, Brown, 1967.

McManus, Walter S.; William Gould and Finis Welch. "Earnings of Hispanic Men: The Role of English Language Proficiency," *Journal of Labor Economics*, 1 (2) (April 1983): 101-30.

McNeill, William H. *Plagues and Peoples* New York: Anchor Books, Doubleday, 1976. (page cited is from 1989 Anchor edition.)

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

Schelling, Thomas. *The Strategy of Conflict*. Cambridge: Harvard University Press, 1960.

------------------. *Micromotives and Macrobehavior*. New York: Norton, 1978.

Sowell, Thomas *Race and Culture: A World View*. New York: Basic Books, 1994.

Weitzman, Martin. "On Diversity," *Quarterly Journal of Economics* 107,2 (May 1992): 363-405.

Table 1

Variable	Obs	Mean	Std. Dev.	Min	Max
cntypct	70782	5.368349	9.614486	.00139	55.11163
age	70782	38.80698	18.72833	6	90
immig	70782	4.777147	2.72701	1	10
fluent	70782	.6806391	.4662323	0	1
Spanish	70782	.4777768	.4995094	0	1
Chinese	70782	.0658077	.2479474	0	1

Table 2

-> fluent= 0

Variable	Obs	Mean	Std. Dev.	Min	Max
cntypct	22605	8.329804	11.58854	.00278	55.11163

-> fluent= 1

Variable	Obs	Mean	Std. Dev.	Min	Max
cntypct	48177	3.978813	8.17041	.00139	55.11163

-> English= 1

Variable	Obs	Mean	Std. Dev.	Min	Max
cntypct	29055	3.573125	7.791676	.00139	55.11163

-> English= 2

Variable	Obs	Mean	Std. Dev.	Min	Max
cntypct	19122	4.595236	8.678413	.00278	55.11163

-> English= 3

Variable	Obs	Mean	Std. Dev.	Min	Max
cntypct	15007	7.15361	10.81885	.00278	55.11163

-> English= 4

Variable	Obs	Mean	Std. Dev.	Min	Max
cntypct	7598	10.65293	12.65892	.010307	55.11163

English	Freq.	Percent	Cum.
1	29055	41.05	41.05
2	19122	27.02	68.06
3	15007	21.20	89.27
4	7598	10.73	100.00
Total	70782	100.00	

Table 3
Logit Estimates: Dependent variable=FLUENT

```
Number of obs  =   65433
chi2(3)        =11981.43
Prob > chi2    = 0.0000
Log Likelihood = -35254.199
Pseudo R2      = 0.1452
```

fluent	Coef.	Std. Err.	z	P>|z|
cntypct	-.0561424	.0009749	-57.589	0.000
age	-.0474538	.0006692	-70.908	0.000
immig	.3953262	.0046764	84.536	0.000
_cons	1.12617	.0225604	49.918	0.000

$\partial \text{ Prob } / \partial x :$

	coeff	$\partial p/\partial x$	std dev	$\partial p/\partial x * std$
cntypct	-0.056	-0.0126	9.61	-0.12
age	-0.047	-0.0103	18.72	-0.19
immig	0.395	0.0860	2.72	0.23

Table 4

Ordered Logit Estimates

```
 Pseudo R2      = 0.0932
 Number of obs =  70782
 chi2(3)        =16984.20
 Prob > chi2   = 0.0000
 Log Likelihood = -82638.876                        Pseudo R2      = 0.0932
```

English	Coef.	Std. Err.	z	P>|z|
cntypct	.0531609	.0007839	67.812	0.000
age	.0479031	.0005166	92.723	0.000
immig	-.3839332	.0036023	-106.580	0.000
_cut1	-.1755134	.0178422	(Ancillary parameters)	
_cut2	1.156458	.0182948		
_cut3	2.77189	.021516		

Prob English = 'not at all'   is   .11 .

∂ (Prob English = not at all) / ∂ (cntypct) = .0052

{∂ (Prob English = not at all) / ∂ (cntypct)} {σ cntypct} = .05 .

Table 5
Spanish SPEAKING SUB-SAMPLE

Logit Estimates   Dependent variable = FLUENT

Number of obs = 33818
chi2(3)   =7139.52
Prob > chi2 = 0.0000
Log Likelihood = -19682.73
Pseudo R2   = 0.1535

fluent	Coef.	Std. Err.	z	P>|z|
cntypct	-.0274512	.0010454	-26.260	0.000
age	-.0640388	.0010498	-61.002	0.000
immig	.4732966	.0069805	67.802	0.000
_cons	.5994525	.0300211	19.968	0.000

Table 6
Chinese SPEAKING SUB-SAMPLE

Logit Estimates   Dependent variable = FLUENT

Number of obs =   4658
chi2(3)   =1204.38
Prob > chi2 = 0.0000
Log Likelihood = -2390.5217
Pseudo R2   = 0.2012

fluent	Coef.	Std. Err.	z	P>|z|
cntypct	-.1680963	.0143737	-11.695	0.000
age	-.0681966	.0026084	-26.145	0.000
immig	.3453852	.0175697	19.658	0.000
_cons	2.3279	.0988421	23.552	0.000