Volume Title: Issues in the Economics of Immigration

Volume Author/Editor: George J. Borjas, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-06631-2

Volume URL: http://www.nber.org/books/borj00-1

Publication Date: January 2000

Chapter Title: Diversity and Immigration

Chapter Author: Edward P. Lazear

Chapter URL: http://www.nber.org/chapters/c6054

Chapter pages in book: (p. 117 - 142)

3 Diversity and Immigration

Edward P. Lazear

A growing number of studies are attempting to document the effect of immigration on wages of native-born Americans.[1] The emphasis has been on a corollary of standard trade theory. The idea is that the immigrant is paid his marginal product. The inframarginal returns are captured by the complementary factors of production, in this case, natives, who own the capital and complementary labor. The focus on wage effects of immigration is a natural consequence.

Most proponents of immigration, however, argue for the diversity value that immigration confers on the United States. The stew tastes better when the ingredients are varied. The notion that the whole is greater than the sum of the parts derives from interactions between factors that somehow add to creativity or other components of output not captured by the standard production function.

There is something to this argument. It would be surprising to find large gains from immigration associated with bringing in more skilled or unskilled workers. Skill is easily arbitraged by new native-born entrants to the labor market. The limit on the difference between the gains to bringing

Edward P. Lazear is the Jack Steele Parker Professor of Economics and Human Resources Management in the Graduate School of Business at Stanford University. He is also senior fellow at the Hoover Institution and a research associate of the National Bureau of Economic Research.

This research was supported in part by the National Science Foundation. This final version incorporates excellent comments from the discussant, Daniel Hamermesh, and from the author's research assistant, Muhamet Yildiz.

1. Borjas (1994) points out that the gains from immigration accrue to the native population precisely when wages are depressed by the entry of immigrants. Studies of the effects of immigrants on natives' wages include LaLonde and Topel (1991) and Card (1990). The studies generally find small, if any, effects of immigration on the wages of natives.

in a skilled versus an unskilled immigrant is the cost of producing a skilled worker from an unskilled one domestically.

In the traditional model, the gain from immigration results from increases in the population, which enhances the value of capital or other factors owned by natives. There is nothing special about immigration. The argument in favor of immigration is identical to arguing that society benefits when everyone has more children because the child will only capture his marginal product. Inframarginal returns flow to capital and other labor owned by his parents or their contemporaries. Selecting the skill level of immigrants is equivalent to determining whether we want more children who will grow up to be skilled workers or unskilled workers.[2]

Even if constant returns to scale prevails, and even if each immigrant brings with him a proportionate amount of new capital, land is fixed and owned by the native population. Any population increase, native born or immigrant driven, causes the value of land to rise, benefiting the native population.

Fans of immigration might claim that this misses the point. It is possible to argue that gains from immigration derive from having a wider economic "gene pool." With less "inbreeding," our ideas may be better and more creative and we are less likely to exacerbate our mistakes. Although it is politically correct to accept the view that diversity provides benefit, there are few studies to document the magnitude of the gains or even that such gains exist.[3]

Diversity surely carries its costs. Because individuals from different cultures have a more difficult time communicating with one another, diversity reduces trade, at least initially. Lazear (forthcoming) analyzed the effects of diversity on trade reduction but ignored any gains to diversity, per se. If the value of diversity is sufficiently large, then perhaps some of the arguments against a heterogeneous workforce could be mitigated or reversed.

The analysis that follows attempts to take the diversity argument seriously. In some sense, it seems a reductio ad absurdum, both at the theoretical and empirical levels. The conclusion is not that certain countries should be favored because of their contribution to diversity, but rather that the current policy, which has the effect of favoring certain countries, does not enhance diversity. An alternative policy that leads to more balanced immigration would further diversity.

2. There is an additional factor. Since parents may care about the well-being of their children more than they do about the well-being of an anonymous immigrant, the wages of the child may enter into the calculation of happiness for the native parent population.

3. See O'Reilly, Williams, and Barsade (1998), who find that the gains from diversity are in fact negative. Because diversity creates conflict, any creativity gains are swamped by those associated with the conflict itself.

Weitzman (1992) models biodiversity. He finds that society does not subsidize the right species in maximizing biodiversity. Other factors, including the "cuteness" of the animal in question, are considerations.

The theoretical analysis builds on the idea that the gains from diversity are greatest when groups have information sets that are disjoint,[4] that are relevant to one another, and that can be learned by the other group at low cost.[5] A more formal model will be presented below, but the intuition can be stated verbally.

First, the diversity gains are greatest when individuals have different information. If information sets are completely disjoint, then members of group A can learn a great deal from group B that they do not already know. If information sets are completely overlapping, then the two groups do not contribute much to each other's knowledge.

Second, the information possessed by the other group must be relevant. For example, the knowledge that an auto mechanic has is quite different from that held by an economist. The information sets are quite distinct and thereby meet the disjointness criterion. But they are not relevant to one another. Knowing how to repair the differential on a 1963 Buick is unlikely to help an economist analyze wage differentials.

Third, even if information sets are disjoint and relevant, they are useless unless they can be understood by the other group. For example, it might be better to express a particular thought in French than in English, but in order for English speakers to get the benefit of this improvement, they must be able to understand French themselves. If it were prohibitively costly to learn the language or obtain the information possessed by the other group, then disjointedness and relevance would have no value.

Diversity is modeled and applied to analyze the choice a country makes about the identity of immigrants. Data from the 1990 census are used to estimate the parameters of the model. The findings are as follows:

1. The current U.S. immigrant flow is inconsistent with diversity. To obtain gains from diversity, it would be necessary to institute a selective immigration policy that eliminates relative-based preferences for immigrants and replaces them with a much more targeted approach. Current American residents may have preferences for their own relatives, per se, but the diversity argument for immigration does not bolster their claims.

2. Ironically, a preference for diversity does not imply a diverse population. When trade with unlike individuals is more valuable than trade with like individuals, the initial population may prefer a homogeneous population of the opposite type. Sale of immigration slots or other transfers may be able to induce the initial population to prefer a heterogeneous population.

3. Groups differ greatly in communication propensity, disjointness, and relevance, the three criteria by which a diverse population can be judged.

4. Hong and Page (1997) focus on the gains from diversity that come about when different agents, each of whom possesses limited ability, work collectively.

5. An informal presentation of these ideas is put forth in Lazear (1998, 310–15).

The current group of immigrants does not do well by any of these criteria. It is possible to select immigrants on the basis of characteristics that would enhance diversity and be consistent with the preferences of the majority of the initial population.

4. Education is an important characteristic, both on the basis of relevance and for communication. As such, an immigration policy that fails to ration slots by price while ignoring the education of immigrants is unlikely to further welfare-enhancing diversity.

5. Immigration policy, more than the underlying characteristics of the countries from which the immigrants are drawn, determines the quality of immigrants observed in the United States. Because the filters are different across groups, immigrants from Japan have lower average levels of education than immigrants from Northern African countries, which is inconsistent with differences in average levels of education in the countries themselves.

6. Balanced immigration, which increases the speed of assimilation, also raises gains from diversity.

3.1 The Model

Let us suppose that there are two groups, A and B. The members of group A have knowledge that spans an interval A_0 to A_1, while the members of group B have knowledge that spans an interval B_0 to B_1. The intervals may be overlapping and the ordering is not important. It is the size of the interval and its overlap that is most important. For simplicity, we reduce knowledge to a scalar variable, x. This is shown in figure 3.1. For example, suppose that the information in question relates to literatures. Then A knows all of the papers on interval A, B knows all of the papers on interval B, and they both know papers on interval AB. As in Lazear (forthcoming), the model is one of random encounter. An individual can encounter one individual per period. This individual is either an A or a B. Initially assume that individuals encounter others based on their proportions in the population—that is; there is no segregation of groups. When an A encounters another A, he can trade with each A receiving surplus equal to

$$\text{Surplus to each A} = A_1 - A_0. \tag{1}$$

Trade with another A can yield surplus because two heads or bodies may be better than one, even when they have the same skills or information. For example, it might be impossible for one person to push a stalled car, but two individuals can complete the task.

When a B encounters another B, she can trade with the B, and each B receives surplus equal to

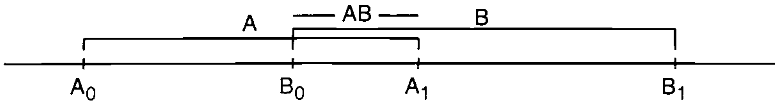

Fig. 3.1 Knowledge

$$\text{(2)} \qquad \text{Surplus to each B} = B_1 - B_0.$$

Trade between A's and B's may create more or less surplus than trade between homogeneous individuals. The surplus could be greater in a heterogeneous pair than in a homogeneous one because the information sets do not overlap completely.[6] In this situation, A learns B's information, which enhances the value of the trade. If B's information were as valuable to A as another A's information, and vice versa, then trade with a B would yield each trader $B_1 - A_0$ of surplus, which exceeds the value of trade between two homogeneous individuals. The less overlap in information sets, the better. This illustrates that trade is enhanced when the information sets of trading parties are disjoint.

Trade between A's and B's might create less surplus than trade between homogeneous pairs. If the information that B's possess is irrelevant to A's activity, then an encounter with a B would not be valuable to an A. Define θ as a relevance parameter. When $\theta = 1$, everything that B knows is relevant to A. Then a trade between an A and a B yields surplus $B_1 - A_0$ to each trading party. When $\theta = 0$, nothing that B knows is relevant to A. It is possible to define θ as unidirectional. Just because B's information is relevant to A does not mean that A's information is relevant to B. Allowing group-specific relevance parameters would add some realism, but the possibility is ignored to conserve on notation.[7] In general, then, surplus between an A and a B who can communicate with one another is given by

$$\text{(3)} \qquad \text{Surplus to each party from diversified trade} = \theta(B_1 - A_0).$$

Third, even if B's information is different and relevant to A, B and A must be able to communicate in order to share the information.[8] If A and B speak different languages, then either A must learn to speak B or B must learn to speak A (or both) in order to share information. Thus, dis-

6. It is assumed that B_0 is never greater than A, to simplify the algebra. This is inessential to any of the results.

7. Logic places a lower bound on θ. If communication were not an issue, then the proportion of B's information that overlaps with A's information must be relevant in order to be consistent with the notion that trade between two A's yields value $A_1 - A_0$ to each party. Since the overlap is $A_1 - B_0$, the lower bound on θ is $(A_1 - B_0) / (B_1 - A_0)$.

8. This is related to the "committee" problem. A large committee possesses more information than any of its individual members, but as the committee gets larger, it becomes impossible to communicate the information. One individual's words drown out another's. This is discussed in Lazear (1998, especially ch. 12).

jointness, relevance, and communication costs determine the value of diversity.

The costliness of communication is modeled by assuming that there is a cost, k_i, for individual i to learn the other language. The distribution function of k is $G_A(k)$ among the A's, and $G_B(k)$ among the B's. It is possible that G_A and G_B are identical, but it is also possible that it is easier for B's to learn A than for A's to learn B, or the converse.

It is now possible to determine how many individuals will learn A and how many will learn B. This depends on the probability of encountering an individual from the opposite group, on the costs of learning, and on the proportion of the opposite group that is bilingual. There is no need for English speakers to learn Spanish if all Spanish speakers also speak English.

Assume only one period. Given that p of the population is A and $1-p$ is B, the expected surplus to a monolingual A is then

(4) Expected surplus to monolingual A

$$= p(A_1 - A_0) + (1 - p)G_B^* \theta(B_1 - A_0),$$

G_B^* is the equilibrium proportion of B's who can speak A. The first term on the right-hand side is the probability of meeting an A times the surplus associated with meeting an A. The second term on the right-hand side reflects the probability of meeting a B times the surplus associated with meeting a B who can speak A, which depends on disjointness and relevance.

Further, the expected surplus to a bilingual A is

(5) Expected surplus to bilingual A

$$= p(A_1 - A_0) + (1 - p)\theta(B_1 - A_0) - k_i,$$

because all encounters with B's result in trade. Note that the cost of learning B is subtracted from the gains from trade.

Analogously, the expected surplus to a monolingual B is

(6A) Expected surplus to monolingual B

$$= (1 - p)(B_1 - B_0) + pG_A^* \theta(B_1 - A_0),$$

and to a bilingual B,

(6B) Expected surplus to bilingual B

$$= (1 - p)(B_1 - B_0) + p\theta(B_1 - A_0) - k_i.$$

3.2 Gains from Diversity

Since A's are in the majority, and since most A's, as an empirical matter, are monolingual, at least in the current stock of Americans, let us consider whether a monolingual A prefers to meet an A or a B. This gets to the heart of the diversity issue. Normalize $A_1 - A_0$ to be equal to 1. Assume, initially, that $B_1 - B_0$ is equal to $A_1 - A_0$ so that A's and B's receive the same value from trading among themselves.

The difference between meeting a B and an A to a monolingual A is given by

Difference in value of meeting a B instead of an A

$$= \theta G_B^*(B_1 - A_0) - (A_1 - A_0) \tag{7}$$

$$= \theta G_B^*(2 - \Delta) - 1,$$

where Δ is defined as $A_1 - B_0$ and is a measure of overlap, which is the complement of disjointness.[9]

The intuitive statements made earlier come directly from differentiating equation (7). First note that

$$\partial/\partial\Delta = -\theta G_B^* + \theta(2 - \Delta)\partial G_B^*/\partial\Delta.$$

Both terms are negative because $\partial G_B^*/\partial\Delta$ is negative (from eq. [9] below). As overlap decreases, for example, disjointness increases, and so does the gain to diversity. When A's can learn more from B's, they are more anxious to encounter B's.

Second,

$$\partial/\partial\theta = G_B^*(2 - \Delta) + \theta(2 - \Delta)\partial G_B^*/\partial\theta.$$

Both terms are positive because $\partial G_B^*/\partial\theta$ is positive (from eq. [9] below). As relevance rises, so does the advantage of meeting a B over an A.

Third,

$$\partial/\partial G_B^* = \theta(2 - \Delta) \geq 0.$$

When more B's can communicate with A's, the gain to a monolingual A from meeting a B is greater. The cost of learning A is a key determinant of G_B^*, so as more B's find it cheap to learn A, group A's gains from diversity rise.[10]

9. If A's and B's were identical and all B's spoke A, then θ would equal 1, Δ would equal 1, and this expression would be equal to zero.

10. Differentiating with respect to G_B^* should be interpreted as a change in the underlying costs of learning a language, which shows up as a change in the equilibrium number of individuals who are bilingual.

Now, the names A and B are arbitrary, except that A's have been defined to be the majority by declaring that $p > 1/2$. But nothing in the above derivation has relied on the fact that $p > 1/2$. Thus, all statements that relate to A's also relate to B's. Specifically, B's prefer to interact with A's when Δ is low. That is, group A's information is disjoint from group B's information in the sense that A's know much that B's do not know. Further, B's preference for interaction with A rises with θ, the relevance of A's information. Finally, as G_A^*, the proportion of A's who can speak B, rises, B's preference for A's rises. Again, as the cost to A's of learning B declines, the B's gain from diversity rises.

3.3 Parochialism

It is useful to consider the conditions under which B's learn A and vice versa. When are groups cosmopolitan and when are they parochial? First, let us consider when a member of the minority will learn A. Taking the difference of equations (6) and (5), the gain to a B from becoming bilingual is

$$(8) \quad \text{Gain to a B from learning A} = p\theta(1 - G_A^*)(B_1 - A_0) - k_i,$$

so that the proportion of B's who learn A are those for whom the right-hand side is positive. Since the distribution function among B's of k_i is $G_B(k_i)$, the proportion of B's who learn A is

$$(9) \quad \text{Proportion of B's who learn A} = G_B[p\theta(1 - G_A^*)(B_1 - A_0)].$$

Analogously,

$$(10) \quad \text{Proportion of A's who learn B} = G_A[(1 - p)\theta(1 - G_B^*)(B_1 - A_0)].$$

By differentiating equations (9) and (10), the following results obtain under general conditions.[11]

First, an increase in p, the proportion of A's in the population, raises the proportion of B's who learn A. This is the primary result of Lazear (forthcoming). Because there are more A's in the population, being able to speak A allows B to trade with more individuals, which is particularly important when B is a small minority and A is a large majority.

Second, the larger is θ, the greater is the proportion of B's who learn A. When A's knowledge is relevant, it pays for B to learn A.

Third, the smaller is G_A^*, the more likely is B to learn A. A small value of G_A^* means that few A's speak B. Thus, the only way for B to trade with A is for B to learn A.

11. The necessary condition is that $p(1 - p)\,\Delta^2\theta^2 G_A G_B < 1$, which is certain to hold as p goes to zero or one.

Finally, the proportion of B's who learn A increases in $B_1 - A_0$, which equals $2 - \Delta$. As the region of information overlap falls—that is, as disjointness rise—so does the value of trading with an A. B types are more likely to learn A when the overlap in information is small.

The population becomes less parochial when disjointness of information is large, when the other group's information is relevant, and when the cost of communication is low.

3.4 Diversity and the Choice of Immigrant Type

It is now possible to consider the primary question. Is there a diversity case for immigration? Let us start by determining which type of individual a country would like to have immigrate.

Initially, there are four types of people in the country. They are monolingual A's, bilingual A's, monolingual B's, and bilingual B's. We assume an egalitarian social welfare function that treats each individual equally and maximizes the sum of surplus across all individuals.

Initially, there are α A's and $(1 - \alpha)$ B's. The equilibrium population will have p A's and $(1 - p)$ B's. The goal is to choose p so as to maximize the welfare of the initial population. Utility of immigrants is ignored.[12]

The utility of an A is given by

$$(11)\quad \begin{aligned} &\text{a. Utility of monolingual A} = p(A_1 - A_0) + (1 - p)G_B^*\theta(B_1 - A_0), \\ &\text{b. Utility of bilingual A} = p(A_1 - A_0) + (1 - p)\theta(B_1 - A_0) - k_i, \end{aligned}$$

and that of a B is given by

$$(12)\quad \begin{aligned} &\text{a. Utility of monolingual B} = (1 - p)(B_1 - B_0) + pG_A^*\theta(B_1 - A_0), \\ &\text{b. Utility of bilingual B} = (1 - p)(B_1 - B_0) + p\theta(B_1 - A_0) - k_i. \end{aligned}$$

Since there are α A's and $(1 - \alpha)$ B's, the expected utility of the initial population as a function of p, the equilibrium proportions, is

$$(13)\quad \begin{aligned} EU(p) = {} & \alpha\{[p(A_1 - A_0) + (1 - p)G_B^*\theta(B_1 - A_0)](1 - G_A^*) \\ & + [p(A_1 - A_0) + \theta(1 - p)(B_1 - A_0) - \bar{k}_A]G_A^*\} \\ & + (1 - \alpha)\{[(1 - p)(B_1 - B_0) + pG_A^*\theta(B_1 - A_0)](1 - G_B^*) \\ & + [(1 - p)(B_1 - B_0) + p\theta(B_1 - A_0) - \bar{k}_B]G_B^*\}, \end{aligned}$$

12. This is not unreasonable since immigrants who come voluntarily are at least made better off by immigration. Further, since they can choose among many countries, competition will induce them to go to the country that provides the best ratio for them, other things constant.

where G^*_A and G^*_B refer to the equilibrium proportions of A's and B's who learn the other group's language where $\bar{k}_A$ and $\bar{k}_B$ are the conditional expectations of k_i, the average cost of learning the language for A's and B's, respectively, given that they learn the other language. The four terms reflect the utilities of monolingual A's, bilingual A's, monolingual B's, and bilingual B's, weighted by their proportions in population.[13]

To find the optimum p, it is necessary to differentiate equation (13) with respect to p. Intuition is gained, however, by considering some specific cases. First, suppose that it is too costly for either A's or B's to learn the other group's language. Then, $G^*_A = G^*_B = 0$, and equation (13) becomes

$$\alpha p(A_1 - A_0) + (1 - \alpha)(1 - p)(B_1 - B_0).$$

Differentiating with respect to p yields

$$\frac{\partial}{\partial p} = \alpha(A_1 - A_0) - (1 - \alpha)(B_1 - B_0),$$

which is positive as long as $\alpha/(1 - \alpha) > (B_1 - B_0)/(A_1 - A_0)$. For $B_1 - B_0 \leq A_1 - A_0$ this is guaranteed because $\alpha > 1/2$. Then $\partial/\partial p$ is always positive, which means that the optimum level of p is 1. If society were to seek immigrants, it would want those who mimic the majority. Under these circumstances, there is no diversity case for immigration. In fact, the reverse is true. Immigration is valuable, but it is because immigration of majority types would increase homogeneity in society. Homogeneity is the desired outcome when individuals cannot trade with members of the opposite group. Because the A's cannot speak B and the B's cannot speak A, it is best to have only one type of individual. Since A's are initially the majority, welfare maximization implies admitting only A's. Under these circumstances, B's , who cannot communicate with A's, will push for more B's. A consequence is that quite divergent views about immigration policy are likely to result.

The divergence in views is, in part, a result of linearity built into the model. There are no diminishing returns to A's from getting more A's. The gains from trade are independent of the number of individuals who have the relevant skills. If diminishing returns were allowed, corner solutions would be less likely. In particular, it is reasonable that a small group might derive more from an additional member with whom it could trade than a large group. Being alone is quite different from having one friend. The one-hundred-and-first friend makes less difference to one's life than the first. If so, then arguments in favor of the majority type to the exclusion of the minority are weakened. Indeed, in an economy where side payments

13. The proportions, G, depend on equilibrium levels, whereas α is the fixed, initial proportion.

are possible (say, by buying immigration rights), all of the nonlinearities would work to affect prices such that an optimal allocation is achieved. This is discussed below.

The strongest case for immigration of minority members can be made when disjointness, relevance, and inexpensive learning prevail. Consider, then, the other extreme, where the cost of learning the other group's language is zero for all individuals. Then, $G_A^* = G_B^* = 1$.[14] Equation (13) is then

$$\alpha[p(A_1 - A_0) + \theta(1 - p)(B_1 - A_0)] + (1 - \alpha)[(1 - p)(B_1 - B_0) + \theta p(B_1 - A_0)].$$

Differentiating with respect to p yields

$$(14) \quad \frac{\partial}{\partial p} = \alpha[(A_1 - A_0) - \theta(B_1 - A_0)] + (1 - \alpha)[\theta(B_1 - A_0) - (B_1 - B_0)].$$

Suppose, for a moment, that A's and B's have equal information. Then equation (14) becomes

$$\frac{\partial}{\partial p} = (2\alpha - 1)[1 - (B_1 - A_0)\theta]$$

or

$$\frac{\partial}{\partial p} = (2\alpha - 1)[1 - (2 - \Delta)\theta].$$

Since $\alpha > 1/2$, the first term is positive. If θ were zero, so that opposite group's information was irrelevant, then $\partial/\partial p$ would always be positive and the optimal p would be 1, as in the case where no one learns the other group's language. Suppose, however, that everything that the other group knows is relevant, so that $\theta = 1$. As long as there is any disjointness at all—that is, as long as overlap is not perfect, so that $\Delta < 1$—the second term is negative, which implies that $\partial/\partial p < 0$ for all values of p. This means that the optimal long-run population would have to be $p = 0$. The initial population would prefer to allow only B's to immigrate because A's get more out of B's than they do out of A's and there are more A's than B's in the initial population.

This produces a surprising implication. Even if there were gains from diversity that outweighed those of trading with one's own type, the implication is that the initial society would want a population of all B's, not a

14. In fact, it is only necessary that one group become bilingual.

diverse population. Homogeneity would be desired, but it would be homogeneity of types in the minority of the initial population. Furthermore, it would be the A's who would prefer this movement to B's; the B's would prefer an opposite movement to A's. The reason is that A's prefer trading with B's, and vice versa. Since there are more A's initially, their preferences win out and the optimum is to admit only B's.

The implication that the society would switch from a majority of A's to unanimity of B's is extreme and is based on the assumption that the initial population can commit to an immigration policy that is maintained in the future. This is unrealistic. If A's were to insist on bringing in only B's, then eventually B's would become the majority. As soon as they outnumbered the A's, they would prefer more A's for the same reason that A's prefer more B's. The majority B's could then institute a policy of admitting only A's, which would hold until A's became the majority, at which point policy would reverse to admit only B's. The equilibrium in a dynamic setting where the majority determines immigration policy is to have a society with half of each type.[15]

The general point is that even when the initial population cares about diversity, it prefers a specific population, not necessarily a diverse population. In some cases, the population preferred by the initial diversity-loving population may be almost completely homogeneous. The immigration rule depends on the size and strength of the various groups, but the message is that the case for diversity as it is generally interpreted is questionable at best on theoretical grounds. Below, it is shown that the case for the current interpretation of diversity fails on empirical grounds as well.[16]

3.5 Immigration and Income

Suppose that we are interested in maximizing GNP or GNP per capita by choosing the number and type of immigrants. Would the solution be the same as that derived above? In general, the answer is no, and the discrepancy between the results in this section and those in the last come about because transfer payments from one group to another have not been allowed.

To maximize GNP per capita, it is sufficient to choose p so as to maximize the net income of the average individual in society, as opposed to the net income of the initial population, which was the criterion expressed in equation (13). To do this, it is necessary to choose p to maximize

15. The situation is much more complicated when there are more than two types because the plurality type prefers a smaller minority group but does not have the power to enforce its desires unilaterally.

16. It is, of course, possible to build in a taste for having a mixed society, just for mixture's sake, but there is no underlying reason for this that comes from the usual arguments having to do with the value of diversity. The enriching value of dealing with other types of people is the basis of the model. To assume a taste for diversity on top of this seems a stretch.

$$
(15)\quad \begin{aligned}
EU(p) = {} & p\{[p(A_1 - A_0) + (1 - p)G_B^* \theta(B_1 - A_0)](1 - G_A^*) \\
& + [p(A_1 - A_0) + \theta(1 - p)(B_1 - A_0) - \bar{k}_A]G_A^*\} \\
& + (1 - p)\{[(1 - p)(B_1 - B_0) + pG_A^* \theta(B_1 - A_0)](1 - G_B^*) \\
& + [(1 - p)(B_1 - B_0) + p\theta(B_1 - A_0) - \bar{k}_B]G_B^*\}.
\end{aligned}
$$

The only difference between equations (13) and (15) is that α is replaced by p in (15) to reflect that we wish to maximize the net income of the average individual in society.[17] First consider the case where learning is free, so that $G_A^* = G_B^* = 1$. Also, assume neutrality so that $A_1 - A_0 = B_1 - B_0 = 1$. Then equation (15) becomes

$$
(16)\quad EU(p) = p^2 + (1 - p)^2 + 2p\theta(1 - p)(2 - \Delta).
$$

Differentiating with respect to p yields

$$
\partial/\partial p = (4p - 2)[1 - (2 - \Delta)\theta],
$$

and again,

$$
\partial^2/\partial^2 p = 4[1 - (2 - \Delta)\theta].
$$

The solution to the first order condition is $p = 1/2$. This is a maximum when $(2 - \Delta)\theta > 1$, or when trading with unlike individuals has more value than trading with like individuals. When diversity has value and is free, the optimal solution for the economy is to choose immigrants to move in the direction of $p = 1/2$. Since $\alpha > 1/2$, this necessarily means that minority immigrants are preferred to majority ones.

This result contrasts with that in the previous section, where the utility of the initial population, rather than overall GNP, was allowed. The reason is this: When trading with unlike individuals is better than trading with like individuals, A's want B's and B's want A's. But since there are more A's than B's in the initial population, maximizing the utility of the initial population pushes the outcome to $p = 0$, with all B's (other than the initial group of A's). This does not maximize GNP. Since it does not, it is inefficient and there is opportunity for trade. One way to deal with the discrepancy is to sell immigration slots, discussed below.

It is possible that trading with unlike individuals has less value than trading with like individuals. Then, the second-order condition implies a minimum. Because it was assumed that A's and B's have the same amount of knowledge—that is, $A_1 - A_0 = B_1 - B_0 = 1$—GNP is maximized by choosing either $p = 0$ or $p = 1$. The tie is broken when one group has more information than another. Suppose, for example, that A's are more

17. Additionally, k changes because the group of individuals that decides to learn the other language varies with the proportion of each type in the overall population.

educated than B's and that education is positively correlated with information and the value of trade. Then this would tip the balance in favor of A's. The value of $EU(p)$ in equation (16) would be maximized at $p = 1$. Allowing $A_1 - A_0$ to differ from $B_1 - B_0$ turns equation (16) into

$$EU(p) = p[p(A_1 - A_0) + (1 - p)(2 - \Delta)] + (1 - p)[(1 - p)(B_1 - B_0) + p(2 - \Delta)], \tag{17}$$

with first-order condition,

$$\partial/\partial p = 2p(A_1 - A_0) - 2(1 - p)(B_1 - B_0) + (2 - 4p)(2 - \Delta)\theta. \tag{18}$$

Setting equation (18) equal to zero yields

$$p = \frac{(B_1 - B_0) - (2 - \Delta)\theta}{(A_1 - A_0) + (B_1 - B_0) - 2(2 - \Delta)\theta},$$

which solves for $p = 1/2$ when B's and A's are symmetric. But if B's have less knowledge than A's, then p moves closer to 1. Some diversity may still be desired. For example, if $A_1 - A_0 = 1.5$, $B_1 - B_0 = .5$, $\Delta = .25$, and $\theta = 1$, then the p that maximizes GNP is 5/6. There is a strong bias toward A's, but complete homogeneity is not desirable.

At the other extreme, when no learning occurs so that $G_A^* = G_B^* = 0$, equation (15) becomes

$$EU(p) = p^2(A_1 - A_0) + (1 - p)^2(B_1 - B_0),$$

which is maximized by setting $p = 1$ if A's know more than B's and $p = 0$ if B's know more than A's. The society should be completely homogeneous because no trade takes place between unlike individuals.[18]

The conclusion of this section is that without transfer payments, the initial population would actually choose immigrants of the opposite type were diversity important. Allowing transfer payments from the new immigrants to the native-born population generally produces an interior solution, but one that favors the group with the most information and skill.

3.6 Selling Immigration Slots

It has been shown that diversity-enhancing immigration may be opposed by the weighted average individual in the initial population[19] even when it would increase overall GNP. Whenever this occurs, there is room for trade. But the ability to buy out the initial population depends on the

18. A solution is to have the group that gets the most out of being in the country "buy out" the other group.

19. It is also true that it will be opposed by the median voter since A's are the majority.

number of immigrants that a country can attract and on the population size that is to be tolerated.

If there is a sufficiently large supply of A's who are willing to immigrate under optimal conditions, and of B's who would immigrate even under the solution preferred by the current population, then it is always possible for B's to bribe the initial group of A's into implementing the GNP-maximizing immigration policy. To see this, denote by R_A^* the surplus that goes to each A under the GNP-maximizing strategy, and by R_B^* the surplus that goes to each B under the GNP-maximizing strategy. Denote by R'_A the surplus that goes to each A under the current-population-preferred solution, and analogously by R'_B for B's. Also, let p^* and p' be the equilibrium proportions under the two regimes. The proof that transfers exist, which make all better off, follows.

Since the average person is better off when GNP is maximized, it must be true that

$$R_A^* p^* + R_B^*(1 - p^*) > R'_A p' + R'_B(1 - p'),$$

or that

$$(R_B^* - R'_B)(1 - p') > (R'_A - R_A^*)p^*.$$

This implies that

$$(19)\quad (R_B^* - R'_B)\frac{B_0 + \dot{B}}{A_0 + B_0 + \dot{A} + \dot{B}} > (R'_A - R_A^*)\frac{A_0}{A_0 + B_0 + \dot{B}},$$

where A_0 and B_0 and are the inital numbers of A's and B's and where $\dot{A}$ and $\dot{B}$ are added to obtain p' in the population. If equation (19) holds, then it must also be true that

$$(R_B^* - R'_B)(B_0 + \dot{B}) > (R'_A - R_A^*)A_0.$$

But this condition says that if each B pays $R_B^* - R_B$, this will compensate every initial A for the loss in moving to the GNP-maximizing solution instead of that chosen by the initial population. However, any B who would immigrate when there are p' A's will certainly move when there are p^* A's. Also, if there is a sufficient number of A's who are willing to immigrate under optimal conditions without compensation, the p^* equilibrium can be achieved. This completes the proof.

3.7 The Empirical Case for Diversity

Theory suggests a way by which having a diverse population can enhance the gains from trade. There are costs of diversity, however, in that communication is hindered when everyone does not speak the same language. Do the gains from diversity outweigh the costs?

To determine whether the argument for diversity has any empirical substance, the 1990 census of the United States (1 percent sample) was used. Data are provided on place of birth, ancestry, English fluency, language spoken at home, and standard variables such as age, education, race, and sex.

It is possible to get a sense of how reasonable the diversity argument is by considering the largest non-English-speaking group in the United States, namely Spanish speakers. Forty-eight percent of those in the 1990 census who were born outside the United States are Spanish speakers. Of those, 55 percent report that they speak English well or very well, which will be defined as fluent. Almost all of the native-born population, which is over 90 percent of the United States, speaks English as its first language. Few in this group speak Spanish. Those who do are, for the most part, children of Spanish-speaking immigrants. Thus, for all intents and purposes, communication does not occur between a native-born American and a Spanish-speaking immigrant unless the immigrant is fluent in English. This means that $G_A^* = 0$ and $G_B^* = .55$.

A necessary condition for the diversity case is that trade between an A and a B results in greater expected surplus than that between an A and an A. Defining A's to be all of the English-speaking, native-born population and B's to be the Spanish-speaking immigrants, it is necessary then that

$$\theta(B_1 - A_0)G_B^* > (A_1 - A_0),$$

or that

(20) $$\theta(B_1 - A_0) > 1/G_B^*.$$

Given that $G_A^* = .55$, equation (20) can be written as

(21) $$\theta(B_1 - A_0) > 1.82.$$

Unless the gains from trading across groups exceeds the normalized 1.82, within-group interaction dominates between-group interaction.

Now, equation (21) is a very difficult condition to meet. To see this, consider a quite extreme situation. Suppose, first, that immigrants have neither more nor less information than do native-born Americans, so that $(B_1 - B_0) = (A_1 - A_0) = 1$. Suppose further that information is almost completely disjoint, with only 10 percent overlap, so that $(B_1 - A_0) = 1.9$. This assumption is very favorable to the diversity case. Also suppose that almost all of the information that each side possesses is relevant to the other side. Specifically, A's find all of the information between A_0 and A_1 relevant and 90 percent of the information between A_1 and B_1 relevant. Then the relevance parameter over the interval $B_1 - A_0$ is

$$\theta = 1(1/1.9) + .9(.9/1.9) = .9526.$$

Under these circumstances,

$$\theta(B_1 - A_0) = (.9526)(1.9) = 1.81.$$

Trade between unlike types is not sufficiently valuable to satisfy condition (20). The expected value of diversity is negative.

Intuitively, since only about half of the immigrants can communicate with the native-born population, each actual trade between an immigrant and a native needs to be worth almost twice as much as that between two natives in order to make the value of diversity positive. It is difficult to imagine that the value of the typical trade between native born and immigrant is almost twice that of the typical trade between two natives. Although possible, the conditions under which diversity pays are very strict.

The main reason for this somewhat negative conclusion is that Spanish-speaking immigrants are not very likely to learn English. Were G^*_B close to 1, the requirement in equation (20) would be much easier to satisfy. This suggests that it is useful to look empirically at how G^*, $B_1 - A_0$, and θ vary across groups. The data in the census files allow us to do this.

3.7.1 Communication

Groups differ greatly in their fluency rates. Table 3.1 reports fluency and education levels among immigrants by region of ancestry. Not surprisingly, immigrants from the British Empire have the highest rate of English fluency. Latin Americans, who constitute the largest group of immigrants, have the lowest fluency rate.

More evidence can be presented on variations in G^* by group. Table 3.2 reports the coefficients on country dummies from a logit that has as its dependent variable FLUENT, a dummy equal to 1 if the respondent reported that he or she spoke English very well or well. The logit is run on the sample of individuals living in the United States in 1990 who were born outside the country and are five years of age or older. Excluded are individuals whose native or only language is English. Thus, Canadians, Australians, and the British are out of the sample. (Of course, by the G^*_B criterion, Canadians are ideal immigrants. They may fall short by the disjointness criterion.) This leaves 147,756 observations.

The right-hand variables include age, years in the United States, and place of birth dummies for the countries listed in table 3.2. These countries are the largest suppliers of immigrants, and they are listed in rank order in table 3.2.

First note that 14 of the 18 coefficients are negative. Because the sample size is so large, all coefficients are estimated with great precision. Statistical significance is not an issue. Relative to the base group, which in this case are those who immigrated from a country not listed in table 3.2, these immigrants are less likely to become fluent in English. This is another

Table 3.1 **Fluency and Education among Immigrants by Region of Ancestry**

Country	Mean
Australia, New Zealand, Canada (N = 2,770)	
Fluent	.98
Education	11.7
Asia (N = 35,338)	
Fluent	.76
Education	11.5
Eastern Europe (N = 11,490)	
Fluent	.86
Education	11.7
Latin America (N = 66,757)	
Fluent	.56
Education	8.7
Middle East (N = 5,495)	
Fluent	.85
Education	12.2
North Africa (N = 574)	
Fluent	.94
Education	14.1
Not specified (N = 14,653)	
Fluent	.76
Education	9.8
Other European (N = 124)	
Fluent	.96
Education	13.2
Pacific Islander (N = 416)	
Fluent	.89
Education	10.9
Sub-Saharan Africa (N = 1,566)	
Fluent	.94
Education	12.9
South Asia (N = 4,762)	
Fluent	.91
Education	13.5
Western Europe (N = 44,031)	
Fluent	.94
Education	11.6
West Indies (N = 5,799)	
Fluent	.94
Education	11.3
U.S. ancestry (not born in United States) (N = 647)	
Fluent	.98
Education	10.5
African American (not born in United States) (N = 1,415)	
Fluent	.98
Education	10.7
Native American (not born in United States) (N = 1,854)	
Fluent	.92
Education	12.4

Table 3.2 **Fluency Logit Analysis**

Country	Coefficient in Logit	Change in Probability
Mexico	−2.287329	−0.497723
Non-Mexico Spanish speaking	−1.412393	−0.307337
China	−1.201402	−0.261425
Philippines	1.385355	0.3014532
Vietnam	−0.975913	−0.212359
Italy	−0.940631	−0.204681
Korea	−0.858997	−0.186918
India	0.8852494	0.1926303
Germany	1.124213	0.2446287
Poland	−0.508278	−0.110601
Russia	−0.631468	−0.137408
Taiwan	−0.183816	−0.039998
Japan	−0.787013	−0.171254
Haiti	−0.428511	−0.093244
Iran	0.4914111	0.1069311
Portugal	−1.424866	−0.310051
Greece	−0.815842	−0.177527
Laos	−1.54405	−0.335985
Other coefficients		
Age	−0.064	−0.014
Years in the United States	0.104	0.022

Note: N = 147,756; log likelihood = −72,679; overall fluency rate = .68.

manifestation of the point made in Lazear (forthcoming). Since these immigrants are from the largest groups, they are the immigrants most likely to encounter individuals with their own backgrounds and therefore are the least likely to learn English.

More important for the purpose here is that there is wide variation across groups. Germans have a predicted fluency probability of .92, whereas Mexicans with the same characteristics have a predicted fluency probability of .37. Older immigrants are less likely to be fluent, consistent with standard human capital predictions. Also, the probability increases by about two percentage points for every additional year that an immigrant is in the United States.

3.7.2 Overlap

The diversity argument relies on the assumption that immigrants have different cultural experiences than native-born Americans and thereby bring new information to the table. But immigrants are not all the same. Some have backgrounds that are very similar to Americans; others are quite different. Although this is difficult to quantify, it is possible to shed some light on the issue by analyzing the ancestry of the American population. These proportions can then be compared to our current flow of immigrants.

Table 3.3 reports the ancestry of a 1/1000 sample of native-born Americans in the 1990 census. The obvious finding from table 3.3 is that about 60 percent of native-born Americans have Western European ancestry. Excluding those who did not specify or who listed U.S. ancestry, the second-largest group consists of African Americans. The third largest group comes from Eastern Europe, followed by Latin American ancestry.

The last column of table 3.3 reports place of birth among the stock of immigrants in 1990. Latin Americans are the largest group, followed by Western Europeans and then by Asians (from East and Southeast Asia). Asian ancestry accounts for less than 1 percent of the native-born population, whereas Latin American ancestry accounts for about 5 percent of native-born Americans. Western Europeans account for 60 percent of the American population. Adding immigrants to the native born drives the Latin American proportion up to around 8 percent and the Asian proportion up to around 3 percent.

Table 3.3 Ancestry among Native-Born Americans in 1990

	N	Frequency in U.S. Population	Frequency of Immigration
By region			
African American	18,382	0.089	0.007
Asia	1,509	0.007	0.178
Australia, New Zealand, Canada	2,262	0.011	0.014
Eastern Europe	12,016	0.058	0.058
Latin America	9,854	0.047	0.338
North Africa	24	0.000	0.003
Native American	6,262	0.030	0.009
Middle East	562	0.003	0.028
Not specified	22,733	0.109	0.074
Other European	381	0.002	0.001
Pacific Islands	285	0.001	0.002
Sub-Saharan Africa	171	0.001	0.008
South Asia	106	0.001	0.024
Stated U.S. ancestry	12,398	0.060	0.003
Western Europe	120,511	0.580	0.223
West Indies	212	0.001	0.029
Total	207,668	1.000	1.000
Selected countries			
China	315	0.002	0.046
Cuba	204	0.001	0.027
Mexico	5,946	0.029	0.197
Philippines	356	0.002	0.041
Vietnam	61	0.000	0.021
Other	200,786	0.967	0.668
Total	207,668	1.000	1.000

The diversity argument suggests that our current immigration policy does not minimize overlap. By the disjointness criterion, the United States admits too many Western Europeans and possibly too many Latin Americans. Asians seem to be the only large group of immigrants that are not already a large part of the American base.

Taken literally, diversity implies that we are accepting the wrong people. For example, underrepresented are North African immigrants. They are the smallest group in the current American population, and there are a significant number of potential immigrants, especially in Egypt, Morocco, and Algeria from which to draw. Indeed, the diversity argument points to a very different immigration policy than the one that is currently in place. Rather than selecting immigrants based on the existence of relatives in the United States, diversity would be served better by doing the opposite. Countries whose residents have the most relatives in the United States are the ones least likely to bring in cultural diversity.

3.7.3 Relevance

The empirical analogue of relevance is somewhat difficult to define. One possibility is that relevance may be related to education. Highly educated immigrants, or at least those with education levels equivalent to those of natives, are more likely to have relevant information than those with much less education. It is unlikely that the details about a particular form of agriculture no longer practiced in the United States are as relevant as information on a new agriculture technique that has been used elsewhere but is not yet practiced in the United States.

It is useful, therefore, to return to table 3.1 and to examine education level by region of origin. It is true, of course, that years of schooling have country-specific meaning. Variations in educational quality and subject matter are likely to be significant across countries. Still, the averages may be instructive.

Somewhat surprisingly, North Africans top the list on average education level. This is almost certainly a result of selective admission. Few and only highly educated North Africans have been successful at obtaining permission to come to the United States. Country-specific evidence is presented in table 3.4, which reports educational attainment (for those no longer in school) by country of origin for large suppliers of immigrants. Immigrants from Mexico have the lowest level of education, and those from India and Taiwan have the highest. Indeed, the highly educated immigrant groups have levels of education that are substantially above the average level among native-born Americans. In 1990, native-born Americans who were not currently enrolled in school and were older than six years old had average levels of education equal to 12.27 years with a standard deviation of 3.08 years.

Again, the differences in education between source countries is as likely

Table 3.4 Mean Levels of Education by Place of Birth

Variable	Observations	Mean	Standard Deviation
Overall	151,888	10.74358	4.76573
Mexico	32,618	7.394307	4.451661
Non-Mexico Spanish speaking	23,599	10.30878	4.50487
China	3,967	10.72044	5.733337
Philippines	7,104	13.17553	3.786291
Vietnam	3,167	10.40101	4.759378
Italy	5,618	9.260413	4.343628
Korea	3,181	12.4967	4.195993
India	3,103	14.83323	4.3834
Germany	6,514	12.33067	3.18601
Poland	3,129	11.10674	4.370777
Russia	2,405	11.58004	4.787556
Taiwan	1,500	14.77033	3.997592
Japan	2,541	13.33806	3.198493
Haiti	1,181	10.46359	4.555955
Iran	1,283	13.86945	4.206208
Portugal	1,549	8.157198	4.717947
Greece	1,575	10.40063	4.338387
Laos	837	6.424134	5.707947

to reflect immigration policy as it is to reflect inherent differences in educational systems or levels. This may be an important point by itself. The characteristics of immigrants in the United States are as likely to reflect the effects of selective immigration policy as they are to reflect the characteristics of the underlying populations from which the individuals are drawn. A policy that is more lenient toward country C than country B will end up with a less qualified pool of immigrants from C than from B, sometimes even when the qualifications of C's are generally higher than those of B's.

The importance of immigration policy in filtering out different groups of immigrants can be seen quite clearly by comparing immigrants from North Africa, whose average education level is 14.1 years, with those from Japan, whose average education level is 13.3 years. The difference observed in the United States between these groups reverses the patterns observed in the native populations and reflects the extreme difficulty of gaining admission to the United States from North Africa.

3.8 Clustering

As mentioned above, the probability of encountering like or unlike individuals is endogenous. In a country that is already as diverse as the United States, it is possible, through geographic mobility, to affect the population with whom trade occurs. Individuals cluster with others of their own type. This is most easily seen by comparing the variable CNTYPCT to the proportion of immigrants in the United States. This variable, discussed in Lazear (forthcoming), measures the proportion of a county's population that is made up of persons who were born in the particular individual's native country.

If immigrants were spread randomly throughout the United States, then the proportion of one's own countrymen encountered would be unrelated to place of residence. Every county would be a microcosm of the United States. For example, 1.7 percent of the people living in the United States in 1990 were born in Mexico. Were they spread randomly throughout the United States, then the average CNTYPCT observed for Mexicans living in the United States would be .017.

Conversely, if Mexican-born immigrants were completely segregated, most counties would have no Mexicans and a few counties would be 100 percent Mexican born. Since CNTYPCT is defined for a specific individual, every Mexican-born immigrant would have a value of CNTYPCT equal to one. That is, every Mexican-born immigrant would reside in a county that consisted entirely of persons born in Mexico.

In fact, the mean value of CNTYPCT among Mexican-born immigrants is .146, much larger than the .017 value that would prevail were Mexican-born immigrants sprinkled randomly throughout the United States. Thus, Mexican immigrants tend to live in more immigrant intensive communities than do natives. The same is true for other large immigrant groups.

The effect of clustering can be examined in another way. A logit identical to that in table 3.2 was run, except that CNTYPCT was included. The coefficient is negative and large. Nonfluent immigrants move to counties with high proportions of individuals from their own countries and they are less likely to learn English. The decision on where to locate is endogenous, and the country coefficients are pushed toward zero when CNTYPCT is included. Taking into account the residential decision reduces the differences between groups because the least-fluent groups are most likely to locate in highly segregated communities. But immigrant groups that segregate pass on fewer of the gains from diversity to the native population.

3.8.1 Diversity Reconsidered

When fluency rates are .55, it is virtually impossible to make a case for diversity. But as the fluency rate rises, the diversity argument makes more sense. Let us consider North African immigrants, whose fluency rate is at .92. Substitution into equation (20) implies that diversity is favored when

$$\theta(B_1 - A_0) > 1.087.$$

Now, suppose that 75 percent of what North Africans know overlaps completely with the native-born American population. Suppose further that the amount of knowledge possessed by natives and by North Africans is the same. Finally, suppose that half of the disjoint information is relevant to native-born Americans. Then, $B_1 = 1.25$, $A_0 = 0$, and

$$\theta = 1(1/1.25) + .5(.25/1.25) = .9,$$

so that

$$\theta(B_1 - A_0) = 1.125,$$

which exceeds the required 1.087. If these assumptions are valid, then diversity, brought about through immigration of North Africans, would be welfare improving to the native population.

The lesson here is that communication between natives and immigrants is the crucial parameter. Unless communication is high, it is virtually impossible to argue in favor of an immigrant group on the basis of gains from diversity. As a practical matter, this means English fluency. Since very few of the native-born American population can be fluent in a large number of other languages, it is necessary that all residents speak a common language.

Of course, it is possible for trade to occur without direct communication. Translators can be used and points of contact between different types of individuals can be minimized. But doing this negates the diversity argument almost by default. An impersonal market, coupled with a few translators, works well to ensure that French wine adorns the tables of American restaurants. But the French vintners need not be U.S. residents for this to occur. The gains from having French vintners teach Californians how to make wine are reaped only when direct communication between the two groups occurs.

Additionally, education and fluency are related. Immigrants, not currently in school, who report that they are fluent in English have average levels of education equal to 11.8 years, whereas those who are not fluent in English have an average education level equal to 7.3 years. Thus, relevance, defined by education level, and communication are likely to be positively related.

Finally, subsequent generations have been ignored. Since virtually all of the children of immigrants are fluent in English, the concerns that were raised in previous sections about English fluency are lessened. On the other hand, children of immigrants who grow up in the United States are less likely to have knowledge and skills that differ from (other) third and subsequent generation Americans. Thus, communication is enhanced when considering children of immigration, but disjointness declines.

3.8.2 Balanced Immigration

The empirical evidence suggest that diversity is enhanced by balanced immigration. Even if one accepts the diversity argument, diversity is useful only when English fluency among immigrants is high. Theory (see Lazear forthcoming) and evidence suggest that individuals who come from countries that make up a small part of the U.S. population are most likely to learn English.

Further, balanced immigration, especially from groups that are not already well represented in the U.S. population, provides the greatest amount of disjointness. If we take the diversity argument seriously, it implies that welfare is enhanced when immigrants come from a large number of underrepresented countries. This suggests that the current policy, which favors relatives of current residents, hinders rather than helps diversity.

Since education is a characteristic that can be screened and selected, there is no obvious reason why countries should be favored or penalized on the basis of the average level of education among their immigrants to the United States. Even though immigrants from Mexico have the lowest average level of educational attainment, nothing prevents the United States from having a policy that favors highly educated Mexican immigrants, if educated immigrants are desired.

Indeed, the lesson learned from this analysis is that current immigration policy is off target if the diversity argument is accepted. Current policy that favors immigrants who have relatives in the United States may have other virtues, but it is likely to grant resident status to those who have significant overlap with the current population, who have low rates of English fluency, and who suffer on the relevance criterion as well.

3.9 Conclusion

A diversity argument can be made for immigration. The desire for diversity is expressed in terms of gains that can be realized by the interaction of individuals who have different backgrounds. Taken literally, the case for diversity is strongest when individuals who differ from the majority confer larger gains from trade on majority members than majority members receive from interacting with their own kind. This argument implies that desirable immigrants come from cultures that are disjoint from current American culture and from cultures that are relevant to Americans. Most important, it is necessary that individuals can communicate with one another. As a practical matter, communication requires a high rate of English fluency among immigrants.

Current immigration policy favors the relatives of U.S. residents. In part, as a result of clustering, this policy has resulted in low fluency rates, which reduces the welfare gains from immigration. Also, because more-educated immigrants are likely to do better on the relevance criterion and because education and English fluency are linked, diversity gains are likely to be positively related to the education levels of the immigrant stock. Related, the results suggest that our immigration policy has resulted in differences in the characteristics of immigrants that reflect the effects of selection as much as they do the underlying characteristics of the populations from which the immigrants are drawn.

The current policy does not lead to an immigrant flow that enhances diversity. Instead, certain countries and cultures are favored at the expense of other countries and cultures. Furthermore, the countries that are the largest suppliers of immigrants are not among the best by the criteria of disjointness, relevance, or communication. A policy that sold immigration slots or one that ranked the specific characteristics of the individual immigrants would be more likely to enhance diversity.

References

Borjas, George J. 1994. The economics of immigration. *Journal of Economic Literature* 32 (4): 1667–717.

Card, David. 1990. The impact of the Mariel boatlift on the Miami labor market. *Industrial and Labor Relations Review* 43 (2): 245–57.

Hong, Lu, and Scott E. Page. 1997. Problem solving by heterogeneous agents. Syracuse University, Syracuse, N.Y.

LaLonde, Robert, and Robert Topel. 1991. Labor market adjustments to increased immigration. In *Immigration, trade and the labor market,* ed. John Abowd and Richard Freeman. Chicago: University of Chicago Press.

Lazear, Edward P. 1998. *Personnel economics for managers.* New York: John Wiley & Sons.

———. Forthcoming. Culture and language. *Journal of Political Economy.*

O'Reilly, Charles, Katherine Williams, and Sigal Barsade. 1998. Group demography and innovation: Does diversity help? In *Research on groups and teams,* vol. 1, ed. E. Mannix and M. Neale. Greenwich, Conn.: JAI Press.

Weitzman, Martin. 1992. On diversity. *Quarterly Journal of Economics* 107 (2): 363–405.