Optimal immigration, assimilation and trade

István Kónya [1]
Boston College[2]

October, 2001

[1] I would like to thank Joseph Altonji, James Anderson, Gadi Barlevy and Kiminori Matsuyama for helpful comments and suggestions.

[2] Correspondence: Economics Department, Boston College. Chestnut Hill, MA 02467. Tel: (617) 552-3690. E-mail: konya@bc.edu

Abstract

The paper develops a model that examines the role of cultural conflict in immigration and immigration policy. Cultural differences lead to frictions between natives and immigrants, unless the latter make a costly investment to assimilate. Since rational agents will assess their assimilation abilities when contemplating migration, the decisions to migrate and assimilate have to be analyzed jointly. The paper's key contribution is to endogenize the migration decision, which highlights the importance of self-selection in the migration process. An important consequence of self-selection is that although there are externalities in the assimilation decision, the equilibrium outcome is efficient for some parameter values. Because of the endogeneity of the migration decision, care must be taken to select the optimal policy instruments. Implementing the first best requires a policy that targets both the size and composition of the migration flow. Therefore, simple policies that only influence one of these are not optimal. In particular, subsidizing assimilation or auctioning immigration permits do not achieve the first best. Instead, a tax scheme that differentiates between assimilating and non-assimilating immigrants can be used to promote efficiency.
Keywords: Culture, immigration policy, optimal immigration.
JEL Classification: D62, F22, Z10

1 Immigration and culture

There has been a marked rise in immigration to the United States in the last two decades, to levels that have not been seen since the beginning of the 20th century. The foreign born proportion of the US population population is now close to 5% and it is rising. Switzerland, Germany and Austria have far more immigrants per capita than the US, while the foreign-born population of France and Britain is roughly the same. Thus migration is clearly a pressing issue both in the traditional immigrant nations of the New World and in Europe.

The economic profession has been quick to analyze some of the consequences of the new migration wave. The emphasis has tended to be on the economic progress of immigrants and their impact on the host country economy. In the labor literature, issues studied include the effect of migration on the wages of low skilled workers and on the welfare state, and the economic assimilation of immigrants in the host country.[1] Researchers in international trade also looked at the economic impact of immigration, and they have explored the connection between foreign trade and immigration.[2] Given the substantial amount of work in this area, we now have a good understanding of these issues. But immigration also has important non-economic effects on immigrants and natives. In practice discussion among non-economists centers on issues that economists have largely ignored, such as cultural frictions and clashes.

Sowell (1996) documents the experience of various immigrant groups throughout history. A recurring theme is that immigrants and natives with different cultures and languages experience frictions in intergroup encounters. Hostility between immigrants and natives can often be traced back to economic reasons, since (as Sowell points out) immigrants tend to hurt vocal native interest groups. But the fact that immigrants can be singled out, based on their different cultural and other characteristics and not on their

[1]Borjas (1999a) provides an overview of this literature. Altonji and Card (1991) and Card and DiNardo (2000) examine the effect of immigrants on local labor markets in the US. The classic references on the economic assimilation of immigrants are Chiswick (1978) and Borjas (1987). Auerbach and Oreopoulos (1999) and Storesletten (2000) use macro-economic models to look at the fiscal impact of immigration.

[2]Trefler (1997) and Venables (1999) look at the connection between migration and international trade in various trade models. Markusen and Zahniser (1997) examine the effect of NAFTA on migration, while Hanson and Slaughter (1999) study the effect of immigration on US production patterns and factor prices.

economic effect, indicates that these cultural differences matter to people.[3] In contemporary Europe, anti-immigrant sentiments are just as strong in low- unemployment Switzerland as they are in high-unemployment Belgium. Clearly, to study the full impact of immigration one has to look beyond the traditional economic factors.

This paper presents a model which examines the role of cultural conflict in immigration and immigration policy. In the model individuals decide whether to migrate into another country. In addition to the migration decision, immigrants can also decide to assimilate – acquire the culture of the host country. In the model this eliminates cultural frictions between them and natives. The crucial feature of the model is the presence of externalities in the assimilation decision, which *might* lead to an inefficient equilibrium outcome[4], both in the size and composition of the migration flow. The latter arises because individuals are heterogenous both in their migration and assimilation costs. This emphasis on the composition of immigration is a novel feature of the model. It leads to implications for immigration policy that are different from what have been suggested in the previous literature.

One of the most important messages of the paper is that the welfare properties of the equilibrium are not uniform across all parameter values. There are structural differences in immigration from developed, middle income and poor countries, and these categories themselves depend on various parameters. A surprising conclusion is that cultural externalities do not necessarily make the equilibrium inefficient. In particular, if the sending country is sufficiently rich, both the level and the composition of immigration are optimal. For less developed source countries this is not the case, so the paper presents the optimal policies that can implement the first best. Because the policymaker has to select not just the size of migration but also those whose assimilation and learning costs are the lowest, the optimal policy has to offer selective incentives to different people. This feature of the model implies that some popular tools of immigration policy do not implement the social optimum, and possibly make matters worse. I analyze the effects of two such policies in detail: assimilation subsidies and restricting immigration levels.

[3] Chinese in Southeast Asia, Jews in medieval and modern Europe and Indians in East Africa suffered discrimination and worse, regardless of their individual economic status. Recurring persecution of Jews in medieval Europe often had no economic reasons, but were based on religious differences and deeply rooted prejudices of the Christian population.

[4] I will look at the welfare properties of the equilibrium from the point of view of a planner who wants to maximize world welfare.

An important special case of the second is selling immigration permits, which is sometimes advocated by economists. I show that in the case of cultural externalities this policy cannot achieve the first best.

The welfare measure used in the paper is the sum of world output. In the model this means that the social planner's objective function includes the welfare of immigrants in addition to the well-being of natives in the host country. Since immigration policy is conducted in the host country, this might require some justification. First, a forward-looking government recognizes that immigrants will be future voters, so it might cater to their interests, especially if they have relatives already in the country. Second, international cooperation on other issues might induce a country to follow policies that take into account immigrant welfare. If, say, the US and Mexico are interested in maintaining long-term cooperation on trade, law enforcement etc, then the US might offer concessions to Mexico in its immigration policy. Finally, it is easy to find the optimal policy for the host country in the model, and I show that in many cases it coincides with the world planner's optimum. When it does not, it is interesting to look at the reasons for the difference, and see if the interests of the social planner and the host country natives can be aligned. I will return to this issue later.

Apart from normative implications the model has interesting positive predictions. It shows, for example, that immigration from culturally very different source countries might have better welfare properties than immigration from similar countries. The model also allows for the possibility that immigrants are isolated from natives in the host country, which is often the case. I show that segregation has ambiguous effects on welfare: it leads to a decrease in the frequency of inefficient encounters, but also hinders assimilation. Finally, the comparative statics results present testable implications, and I provide some preliminary evidence that the model's predictions are born out in US data.

One strand of the relevant literature is work on optimal migration, where authors looked at causes why migration levels might not be Pareto-efficient from the point of view of world welfare. In a book that discusses the possible effects of Eastern European migration to Western Europe (Layard, Blanchard, Dornbusch and Krugman 1992) the authors list many possibilities that are of concern (wage compression, fiscal externalities, market size effects, information and endogenous tastes, and cultural externalities), but they do not develop formal models to study them. Burda and Wyplosz (1992) explore the role of human capital externalities in the context of migration from Eastern to Western Europe, using a Lucas-type growth model. The main difference between their model and the one in this paper is that the composition of immigration is crucial here, while in Burda and Wyplosz (1992) only the size of migration matters. This leads to very different policy implications, as we will see later.

The literature on optimal migration has not explored the role of cultural differences. Two path-breaking articles (Lazear 1995, Lazear 1996), however, look at cultural externalities and assimilation from a partial equilibrium point of view. Lazear models interaction between immigrants and natives as random matching. I depart from his model by explicitly considering the migration decision. Lazear treats the size and composition of an immigrant group as exogenous, which might be the case for some specific groups (most notably, refugees). In general, however, immigrants are self- selected along many dimensions, among them the capability to assimilate. It is reasonable to expect that those who actually migrate are the ones who can learn the new culture easier. Migration and assimilation decisions are not independent, but arise from the same fundamentals, and my model traces both back to these fundamentals. The importance of endogenizing migration is highlighted when we look at policies such as assimilation subsidies to immigrants. In Lazear's partial equilibrium framework assimilation subsidies can be used to implement the first best. When migration is endogenous this is no longer the case, because assimilation subsidies fail to select the immigrants who should come in the first place.

The rest of the paper is organized as follows. In Section 2 I describe the formal model of the migration and assimilation decisions in a two-country framework. In Section 3 I derive the equilibrium conditions, explore the existence and uniqueness of the equilibrium and look at some comparative statics results. Section 4 solves the social planner's problem, and Section 5 discusses the policy implications of the model. In Section 6 I present some evidence in support of the model, using data from the United States. Finally, Section 7 concludes.

2 The model

2.1 Assumptions

In this section I develop a model of migration, assimilation and cultural externalities. A simple and transparent way to model interaction between people is to use a random matching framework, which in this context was first introduced by Lazear (1995). In such a setting agents are paired together according to a random device, and generate some output that they share according to a sharing rule.[5] In contrast to assortative matching, people do not have the ability to search for good pairs, and might be stuck in an inefficient allocation. There is only one match for each agent, and agents derive utility from the "income" (output share) they receive from this match. I assume the utility function is linear, and I normalize utility units such that they are equal to the output units. The qualitative conclusions would remain the same with a more general concave utility function, but the linearity assumption simplifies the analysis greatly.

There are two countries, North and South, with initial populations of 1 and L, respectively. Agents can change locations, but moving is costly. If a person decides to migrate, she has to pay c units of her income. Populations are heterogenous with respect to c, and the initial distribution of the moving costs in each country is given by the c.d.f. $F(c)$, with $c \in (0, K)$. The two countries differ in their culture, which influences the outcome of trades. When two people who have different cultures are matched together, the output they generate is only $\theta < 1$ times the output they would receive if they had the same culture. Also, the South is poorer than the North and a match in the South is worth only $h < 1$ times as much as the same trade would be worth in the North. There are no international matches – all trade takes place within the countries.[6] Finally, I assume that the output from a match is shared equally by the two parties, and I normalize the value of a uni-cultural match in the North to 1 for each trader.

Because the South is poorer, in equilibrium migration will flow from the South to the North. When immigrants arrive at the North, their culture is different from the native one. Thus matches between immigrants and natives

[5] In the rest of the paper I will also use the phrase "trade" to indicate a match.

[6] Although it is possible to relax this assumption, there is no obvious way to incorporate international trade into the model. Since the basic conclusions do not change with such an extension, I use the closed economy model to avoid unnecessary complications.

would be less efficient than a trade within a group. It is possible, however, that an immigrant learns the native culture – that is, she assimilates. An assimilated immigrant can trade efficiently with other immigrants and natives, because she knows both cultures. Thus in the North trades result in an output of 1 to the parties, unless a native and a non-assimilated immigrant meet. In this case they only get θ each. Since non-assimilated immigrants have a Southern culture, the parameter θ measures the cultural distinctiveness between the North and the South. The more similar the two countries are, the more efficient a bi-cultural match in the North is. Finally, note that the South remains homogenous (no immigrants there), but because it is less developed, trades yield $h < 1$ in any Southern match for the two traders.

Immigrants enjoy more efficient native matches when they assimilate. However, learning the native culture is costly, and immigrants have to forgo a fraction τ of their output in order to learn. Thus assimilation involves a time cost, the time required for learning. τ is different for each individual, and its distribution among the *initial* populations in the two countries is given by the c.d.f. $G(\tau)$, with $\tau \in [0, 1]$.[7] It is important to note that the distribution of τ among immigrants is different from the distribution in the general population. This is because the capability to assimilate influences the migration decision, therefore immigrants will be self-selected in their capacity to assimilate. To make inferences about optimal immigration and immigration policy, we have to take this dependence into account.

Let us now turn to the matching technology. A stylized fact in immigration is that people tend to live with others who share their culture. In a book about Vietnamese refugees in the United States, Zhou and Bankston III (1998) document that although the US government made a conscious effort to scatter Vietnamese among natives, after a few years they tended to move into areas with high concentration of their countrymen. In general, immigrants who do not (or only partially) assimilate live in their own neighborhoods, and hence their contact with natives is limited. Assimilated immigrants, on the other hand, tend to move into native areas, because of generally better amenities and job opportunities. To capture these stylized facts, I assume that the Northern population is geographically segregated into two segments. One consists of natives and assimilated immigrants, while the

[7]It would be reasonable to assume that the distribution function G also depends on θ, because it might be easier to learn for an immigrant if she comes from a more similar culture. I will use the simplest formulation that does not allow for such dependence, but my results remain the same in the more general setting.

other includes non-assimilated immigrants. There is more interaction within each group than between them, and the difference depends on the level of segregation in the society. At one extreme, all immigrants are scattered randomly among natives and all matches are equally likely. At the other extreme, non-assimilated immigrants are completely isolated and there is no inter-group match.

Let p be the probability that two non-assimilated immigrants meet, and let q be the probability that a native or an assimilated immigrant meets someone from their group. Also, let the equilibrium number of immigrants be m, and let the number of assimilated immigrants be a (thus the post-migration population of the North is $1 + m$).[8] I assume that p and q are given by the following expressions:

$$p = \alpha \frac{m - a}{1 + m} + 1 - \alpha \tag{1}$$

$$q = \alpha \frac{1 + a}{1 + m} + 1 - \alpha. \tag{2}$$

Thus α measures the level of interaction between the two groups: when $\alpha = 1$ intra- and inter-group matches occur with the same probability, when $\alpha = 0$ there are only intra-group trades. Note that trades must be balanced between groups, and it is easy to check that for the above specifications for p and q this is indeed the case.[9] Finally, I assume that non-assimilated immigrants with a match from the other group meet a native with probability $1/(1 + a)$ and an assimilated immigrant with probability $a/(1 + a)$. Thus non-assimilated immigrants cannot search actively for assimilated ones.

To finish this section, I will maintain the following assumptions about the distribution functions in the paper. First, both F and G are differentiable and have a continuous p.d.f. f and g, respectively. Second, both distribution functions have full support. Third, c and τ are independent. If we associate the assimilation cost with ability, we can give examples suggesting both positive and negative correlations. More able people might be more efficient movers, leading to a positive correlation. But they might have more to loose

[8]It turns out that it is more convenient to work with the overall number of immigrants than with the number of non- assimilated ones.

[9]The number of natives and assimilated immigrants who are matched with the other group is $(1 - q)(1 + a)/(1 + m)$, while the number of non-assimilated immigrants who are matched outside is $(1 - p)(m - a)/(1 + m)$. One can easily check that the two quantities are the same.

from moving, which leads to a negative correlation. Given this ambiguity, it is natural to start with the assumption of independence, which also makes the analysis much easier.

2.2 Individual decisions

Let us now examine how people choose to move and assimilate. In equilibrium there are four different groups of people: Northern natives, assimilated immigrants and non-assimilated immigrants in the North, and Southerners who do not migrate. In equilibrium the four group sizes are 1, a, $m-a$ and $L-m$, respectively. The endogenous variables are m and a, and our goal is to find these values. To start, let us write down the incomes or value functions for each group, given our assumptions on the matching technology and on the efficiency of particular types of matches from the previous section:

$$\begin{aligned} V_N &= q + \theta(1-q) \\ V_a &= (1-\tau)(q+1-q) - c \\ V_n &= p + (1-p)\,\frac{a+\theta}{1+a} - c \\ V_S &= h, \end{aligned}$$

where N, n, a and S stand for Northern natives, non-assimilated immigrants, assimilated immigrants and Southerners. Each term is the value of a particular match type multiplied by the probability of such a match. Note that for non-assimilated immigrants the second term is the expected value of an intergroup trade, since the partner can be a native or an assimilated immigrant in such cases.

Using the definitions of p and q from (1) and (2), we can rewrite the value functions in the following forms:

$$\begin{aligned} V_N &= 1 - \alpha\frac{(1-\theta)(m-a)}{1+m} \\ V_a &= 1 - \tau - c \\ V_n &= 1 - \alpha\frac{1-\theta}{1+m} - c \\ V_S &= h. \end{aligned}$$

Notice that keeping the number of immigrants constant, Northern welfare is increasing in a, the number of assimilated immigrants. This is the case because natives can trade more efficiently with assimilated immigrants, so they benefit when immigrants learn their culture. Immigrants, however, do not take into account this benefit for natives when they decide about learning. This external effect of the learning decision is what we called cultural externalities in the introduction, and now we can start to examine their effect on the migration and assimilation outcomes.

The difficulty to solve the model lies in the fact that the migration decision is influenced by the assimilation choice. That is, individuals in the South can choose among three possibilities, and whether they migrate depends on their capability to assimilate. To overcome this barrier, I will proceed the following way. First, it is evident that for any given m and a, the individual assimilation decisions involve a threshold. Immigrants will assimilate only if their benefit from doing so is greater than their cost. Equating V_n and V_a yields the threshold level for τ:

$$\tau_a = \frac{\alpha(1-\theta)}{1+m}. \tag{3}$$

Thus any immigrant with a lower learning cost will assimilate, and others will not. This *does not* mean, however, that in a given immigrant group some people do not assimilate. As I stressed earlier, the distribution of τ is endogenous, and whether agents with high learning costs choose to leave the South depends on the underlying parameters. For this reason it is meaningless to do comparative statics exercises with τ_a, since we do not know what the cutoff means if the distribution of τ changes with the parameters. In order to say more, we have to examine the migration decision.

We can start with those people who migrate but don't assimilate. Let us indicate the threshold in c for them with c_n. We get this value if we equate V_S and V_n. That is, we have:

$$c_n = 1 - h - \frac{\alpha(1-\theta)}{1+m}. \tag{4}$$

For the people who assimilate, the relevant threshold can be calculated by equating V_S and V_a. If we rearrange this condition, we end up with the following expression:

$$c_a = 1 - h - \tau. \tag{5}$$

Thus we have three threshold levels that together determine the choice of the Southern people. Figure 1 shows the break-down of the Southern

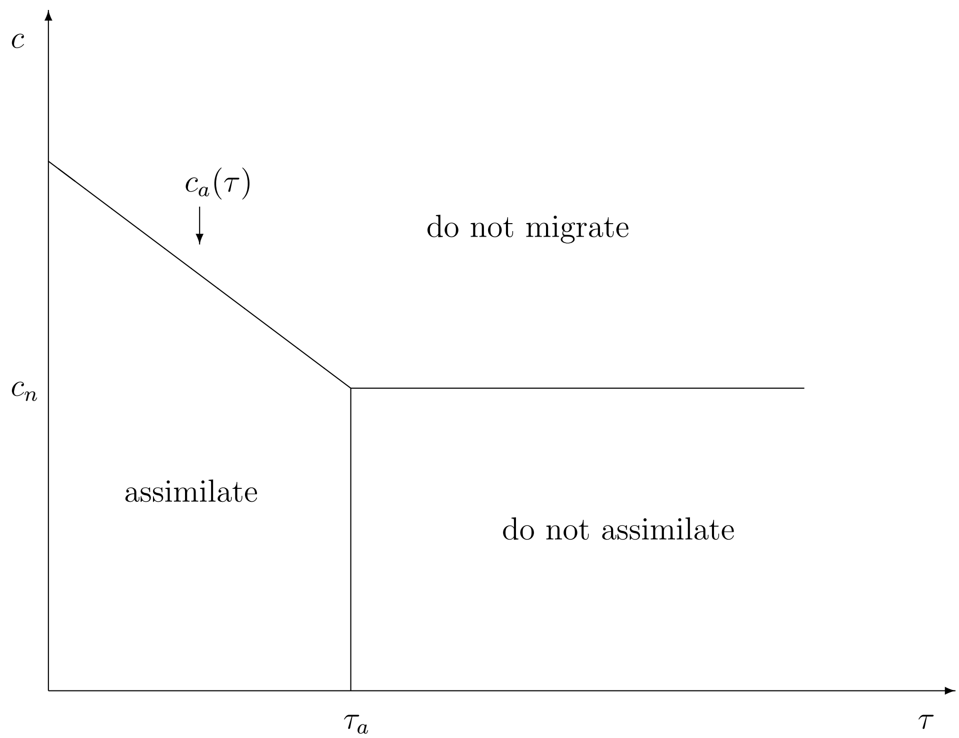

Figure 1: The migration-assimilation decision

population in the (τ, c) space. Those with small enough τ and c migrate and assimilate. People with low c and high τ migrate, but do not assimilate. Finally, the rest of the population stays in the South. Two things are worth noting. First, the location of the c_n line is endogenous, since c_n depends on m. Second, the c_n line can be below zero, in which case non-assimilating migration will not take place. But there are still assimilating immigrants, since the $c_a(\tau)$ line starts above the c_n line at $\tau = 0$ and it cuts the c axes above zero (since $h < 1$). This suggests that there might be two types of equilibria, one with non-assimilating immigration and another without it. Let us now turn to the formal analysis where I show that this is indeed the case.

3 Equilibrium

3.1 Existence and uniqueness

We can start by verifying the two observations from Figure 1. First, if there are non-assimilating immigrants, there also must be assimilating ones. From (4) and (5) we can see that $c_a(\tau) \geq c_n$, because the last term in c_n is just τ_a, which is the upper limit on learning costs for an assimilating immigrant. In other words, since assimilation is a profitable activity, its benefits can be traded off for higher migration costs. That is, low τ individuals will migrate even if their c is above the distribution of c for those who do not assimilate. Second, it is easy to see that migration is always profitable for someone. Since the lower limit of τ and c is 0, for any $h < 1$ very low cost individuals can profit from assimilating migration.

These two results confirm our conjecture that two types of equilibria are possible. Let us call the one with both types of immigrants *interior* ($0 < a < m \leq L$) and the other with only assimilating immigrants *corner* ($0 < a = m \leq L$). If the solution is interior, it has to satisfy the following:

$$m - a = LF(c_n)\left[1 - G(\tau_a)\right] \tag{6}$$

$$a = L \int_0^{\tau_a} F[c_a(\tau)]\, dG(\tau) \tag{7}$$

The equations relate the expected and actual number of migrants of different types and τ_a, c_n and c_a are as defined in (3), (4) and (5). In (6) the right-hand side is the count of people who would not assimilate ($\tau > \tau_a$), but have low enough moving cost to migrate ($c < c_n$). Since c and τ are independent, the mass of such people is just a product of the two probabilities times the total mass of people in the South (L). In (7) we have a complication: c_a depends on τ. Thus first we count people with low moving costs ($c < c_a$) for a given τ, and than "add" them together (integrate) for all $\tau < \tau_a$. Notice that we can eliminate a because it does not enter the terms on the right-hand sides of (6) and (7). This gives us the following condition for m:

$$m = LF(c_n)[1 - G(\tau_a)] + L \int_0^{\tau_a} F[c_a(\tau)]\, dG(\tau). \tag{8}$$

If $a = m$ we need only one equation, since non-assimilating migration is zero. This means that all immigrants have $\tau \leq \tau_a$, possibly with the inequality strict. In other words, we cannot just use (7) to determine a when $a = m$, since the upper limit on assimilation costs among immigrants might be lower than τ_a. To get the appropriate cutoff we have to solve $c_a(\tau) = 0$, which leads to the value of $1 - h$.[10] Thus in the corner case the equilibrium condition is:

$$m = L \int_0^{1-h} F(1 - h - \tau)\, dG(\tau), \tag{9}$$

Since for any $c \leq 0$ we have $F(c) = 0$, we can summarize the equilibrium conditions as follows:

$$m = LF(c_n)[1 - G(\tau_a)] + L \int_0^{\min\{1-h,\tau_a\}} F(1 - h - \tau)\, dG(\tau)$$

$$\text{s.t.} \qquad \tau_a = \frac{\alpha(1-\theta)}{1+m} \qquad \text{and} \qquad c_n = 1 - h - \frac{\alpha(1-\theta)}{1+m}, \tag{10}$$

where for convenience I repeated the equations for the cutoffs (and substituted for c_a). Thus we have to examine (10) to establish the properties of equilibrium.

It is easy to see that an equilibrium exists, as the following proposition shows.

Proposition 1. *For any parameter values* $0 \leq \alpha \leq 1$, $0 < h < 1$, $0 \leq \theta < 1$ *and* $L > 0$ *there is an equilibrium level of migration* $0 \leq m \leq L$.

Proof. The proof is a straightforward application of Brouwer's Fixed Point Theorem. Let us denote the right-hand side of (10) by $\Gamma(m)$, thus the equilibrium condition is to find the fixed point of $\Gamma(m)$ on the set $[0, L]$. This set is compact and convex, while the function $\Gamma(m)$ is continuous (though not necessarily differentiable) and maps $[0, L]$ into itself. Thus the conditions of Brouwer's Theorem are satisfied, and there is an $m \in [0, L]$ such that $m = \Gamma(m)$. □

Thus the existence of an equilibrium is guaranteed, but uniqueness is not. The solution to (9) is unique (the right-hand side is independent of m), thus there cannot be more than one corner equilibrium. Equation (8), however, can be satisfied by more than one m. It is easy to see the intuition behind the possibility of multiple equilibria in this case. Looking at (3) and (4) reveals that τ_a decreases and c_n increases with m. Thus there is a positive

[10] It is easy to check that $c_n < 0$ implies $1 - h < \tau_a$.

feedback in immigration: the more people move, the higher the incentives for non- assimilating immigration. Since those who do not assimilate trade efficiently only with other immigrants, it is better if they are more numerous (and have a higher probability to meet). To see the problem mathematically, let us take the derivative of the right-hand side of (8), $I(m)$:

$$\frac{\partial I(m)}{\partial m} = Lf(c_n)[1 - G(\tau_a)]\frac{\alpha(1-\theta)}{(1+m)^2} > 0.$$

This expression can take any values between zero and infinity, so the function $I(m)$ can cut the 45^0 line more than once. It is possible, however, to rule out some of the possible equilibria on stability grounds. Although the model is static, we can talk about Walrasian or tâtonment stability. In this context the equilibrium is locally stable if $m - I(m)$ is increasing around the equilibrium point or $I'(m) < 1$. If this condition holds independent of m the solution to (8) is unique, because $0 - I(0) < 0$, $1 - I(1) > 0$ and $m - I(m)$ is monotone decreasing.

In the rest of the paper I will restrict attention to the case where the solution to (8) is unique. A sufficient condition for this is given by the following restriction:

Assumption 1. *The distribution of moving costs is not very concentrated. More precisely, for any values of $c \in (0, K)$ the following inequality holds:*

$$f(c) < \frac{1}{\alpha(1-\theta)L}. \tag{11}$$

Note that Assumption 1 is likely to hold when α is small (North is very segregated), θ is close to one (cultural differences are small) or when L is small. Even in the polar case of $\alpha = 1$ and $\theta = 0$ (and $L = 1$), the condition is satisfied when F is uniform on $(0, K)$, where $K > 1$. Thus Assumption 1 is a fairly mild restriction, and is likely to be satisfied if people are heterogenous enough in their moving costs. Finally, even when multiplicity arises, many results below (e.g. comparative statics) continue to hold for stable equilibria.

Even if the solution to (8) is unique, it could still happen that there is an interior and a corner equilibrium for the same parameter values. As the next proposition shows, however, such an outcome is not possible.

Proposition 2. *Given Assumption 1, the solution (m^*, a^*) to the immigration problem is unique. Moreover, $\exists! \bar{h}(\alpha, \theta, L)$ such that if $h < \bar{h}$ the equilibrium is interior and if $h \geq \bar{h}$, $0 < a^* = m^* < 1$. Thus non-assimilating migration only occurs when the South is sufficiently poor.*

Proof. We have already showed that there is a unique solution to both (8) and (9). The only question is whether there is a range of parameter values when both of these m's are equilibrium ones. It is clear that in any interior solution $c_n > 0$. Let us define $\bar{m}$ by the equation $c_n = 0$. That is,

$$\bar{m} = \frac{\alpha(1-\theta)}{1-h} - 1.$$

We know that c_n is increasing in m. This means that we have interior equilibria if the solution to (8) is greater than $\bar{m}$ and a corner solution if $m^* < \bar{m}$. At the cutoff $\bar{m}$ the two types of equilibria should coincide. This defines $\bar{h}$ implicitly with the following equation:

$$\frac{\alpha(1-\theta)}{1-h} - 1 = \int_0^{1-\bar{h}} F(1-\tau-\bar{h})\, dG(\tau), \tag{12}$$

which we get if we substitute $\bar{m}$ into either (8) or (9). We can easily see that there is a unique $\bar{h}$ that satisfies (12): the left hand side is increasing and the right hand side is decreasing in h, moreover l.h.s<r.h.s at $h = 0$ and l.h.s>r.h.s at $h = 1$.

The final step is to show that if $h < \bar{h}$ then $m^* > \bar{m}$ and if $h > \bar{h}$ then $m^* < \bar{m}$. That is, the separation of the equilibria is consistent with the condition on c_n. But this is easy, since the right-hand sides of (8) and (9) are decreasing in h (see the Appendix), so a higher h means a smaller m in both types of equilibria. Thus for $h > \bar{h}$ we have $m < \bar{m}$ and hence $c_n < 0$, and the reverse is true for $h < \bar{h}$. □

Proposition 2 shows us that immigration has a different character for medium income and poor countries.[11] Since the former are relatively well off, migrating is attractive only for a few people. Thus immigrants from middle income countries tend to be those who are mobile *and* assimilate easily. The situation is different for poorer sending countries ($h < \bar{h}$). When h is small, gains from immigration are large even if no assimilation occurs. Therefore in this case many people move even if it is very costly for them to integrate into the North.

[11]Notice that this statement is only true when $\alpha > 0$. When isolation is complete there are no incentives to assimilate, since there is no inefficient interaction with natives. Since my goal is to evaluate the effect of cultural externalities when they are present, I will ignore the case of $\alpha = 0$.

3.2 Comparative statics

We already saw that poorer countries (small h) send more immigrants, which is a very intuitive result. At some point $(\bar{h})$ there is a structural switch, as we move from one equilibrium type to another. We know that non- assimilating immigration occurs if the South is sufficiently poor. We would also like to know if the number of non- assimilating migrants $(m - a)$ increases with h. The answer is affirmative, as the Appendix shows. In this sense, assimilation is a bigger problem when the sending country is poorer.

The second parameter of the model is θ, the cultural similarity of North and South. It is easy to show (see the Appendix) that more cultural similarity (a higher θ) leads to higher immigration in both types of equilibria. Also, from (12) we can see that $\bar{h}_\theta > 0$, so more cultural similarity makes non-assimilating migration feasible for relatively richer countries. This result is interesting because we would expect migration to cause more problems from culturally very different sending countries. This is true if we only look at the direct effect (less efficient matches between natives and non-assimilating immigrants). But there is also an indirect effect that works in the opposite direction and follows from the endogeneity of the migration decision. Since culturally more distinct people are less likely to migrate, they are also less likely to come if they cannot assimilate. Thus when comparing two source countries, North might want to prefer immigration from the less similar one, because all immigrants assimilate. This prediction of the model might explain why US immigrants from the Far East are so successful: they are self-selected in their desire to assimilate. If cultural differences are too big, migration is only profitable for those who can integrate into the host society.

The third parameter is L, the relative size of the source country. Intuitively we expect that a bigger South sends more immigrants, since the distribution of moving and assimilation costs does not depend on L. As the Appendix shows, this intuition is correct, and m depends positively on L. Moreover, the amount of non-assimilating immigration $(m-a)$ also increases with the size of the South in any interior equilibria. These two results are very natural, and arise because in a bigger South there are simply more people to take advantage of any gains from migration. What happens with the cutoff $\bar{h}$ when L changes? It is easy to see from (12) that the cutoff is increasing in L. This result is a consequence of the scale effects in non-assimilating immigration. When more people migrate, assimilation is less attractive because there is a bigger chance that immigrants are matched together. But an increase in L leads to a higher m, which in turn raises the cutoff c_n. Thus for some values of h where the equilibrium was previously a corner, non-assimilating migration now becomes profitable. That is, a larger country not only sends more non-assimilating immigrants in an interior equilibrium, but it is also more likely that such an equilibrium arises.

The last parameter to look at is the level of contact between the two segments of Northern society, α. The Appendix proves that both the level of migration and the level of non-assimilating migration decrease with α, while the level of assimilating migration increases. That is, the more detached the two groups in the North are, the more likely that immigrants will come but do not assimilate. Thus isolation has two effects on native welfare that are of opposite signs. First, less contact between the two groups means fewer inefficient encounters. On the other hand less contact also discourages assimilation and encourages non-assimilating migration, which hurt natives. Moreover, $\bar{h}$ decreases with α (see [12]), thus non-assimilating migration is more likely when the North is highly segregated. As long as there is some contact between the two cultures, an increase in isolation might hurt natives, even in the absence of positive spillovers between groups. I will return to this issue when I discuss possible policies for migration and assimilation. Before that, however, we have to compare the equilibrium with the planner's solution, and this is the task of the next section.

4 Global optimum

In this section I examine what the optimal size and composition of the migration flow would be from the point of view of the social planner who wants to maximize world welfare. I assume that the planner maximizes total output, which is the (population) weighted sum of the four value functions. Given the optimal values of $\hat{m}$ and $\hat{a}$, the planner selects the lowest cost individuals for the two migration types. Thus the optimal decision is still characterized by cutoffs in τ and c. Let us indicate these cutoffs by T_a, C_n and C_a.

As in equilibrium, it is possible that the planner chooses a corner solution, that is $\hat{m} = \hat{a}$. Thus the formal maximization problem can be written as follows:

$$\max_{m,a,T_a,C_n,C_a(\tau)} \left\{ h(L-m) + L[1-G(T_a)] \int_0^{C_n} \left(1 - \alpha\frac{1-\theta}{1+m} - c\right) dF(c) \right.$$
$$\left. + L\int_0^{T_a}\int_0^{C_a(\tau)} (1-\tau-c)\, dF(c)\, dG(\tau) + 1 - \alpha\frac{(1-\theta)(m-a)}{1+m} \right\}$$

$$\begin{aligned} \text{s.t.} \qquad a &= L\int_0^{T_a} F[(C_a(\tau)]\, dG(\tau) \\ m-a &= LF(C_n)\,[1-G(T_a)] \\ m-a &\geq 0. \end{aligned} \tag{13}$$

Notice that the cutoff level for assimilating immigration is defined as a function, since it might (and will) be different for different values of the assimilation cost τ.

It is convenient to rewrite the maximand in a more suitable form. Using the first two constraints and denoting the multiplicators by λ, ν and η, the Lagrangian for the problem can be simplified to:

$$\mathcal{L} = (1-h)m - \frac{2\alpha(1-\theta)(m-a)}{1+m} - L[1-G(T_a)]\int_0^{C_n} c\, dF(c)$$
$$- L\int_0^{T_n}\int_0^{C_a(\tau)} (c+\tau)\, dF(c)\, dG(\tau) + \lambda[L\int_0^{T_a} F(c_a(\tau))\, dG(\tau) - a]$$
$$+ \nu[LF(C_n)(1-G(T_a)) - m + a] + \eta(m-a).$$

To characterize the optimum, let us take the first-order conditions with respect to m, a, T_a, C_n, $C_a(\tau)$. If the optimum is interior ($m > a$ and $\eta = 0$), we can rearrange the FOC's and the constraints and arrive at the following system of equations:

$$\begin{aligned} a &= L\int_0^{T_a} F(C_a(\tau))\, dG(\tau) \\ m-a &= LF(C_n)(1-G(T_a)), \end{aligned} \tag{14}$$

where

$$T_a = \frac{2\alpha(1-\theta)}{1+m}$$
$$C_n = 1 - h - \frac{2\alpha(1-\theta)(1+a)}{(1+m)^2}$$
$$C_a(\tau) = 1 - h + \frac{2\alpha(1-\theta)(m-a)}{(1+m)^2} - \tau.$$

In a corner optimum, where $m = a$, the migration level is given by:

$$m = L\int_0^{1-h} F(1-h-\tau)\, dG(\tau). \tag{15}$$

Although the conditions above are suggestive of the pattern of the optimal solution, there are some technical difficulties involved. The problem is that the system of equations that corresponds to the interior case does not necessarily give a unique solution, and there might be more than one local optima in that case. Fortunately, we can still tell a great deal about the optimum when the solution to the planner's problem is unique,[12] and Proposition 3 summarizes these results.

Proposition 3. *Assume that the solution to the social planner's problem is unique. Then there is an* $0 \leq \hat{h}(\alpha, \theta, L) < 1$ *such that when* $h < \hat{h}$ *the optimal solution is interior. Otherwise non-assimilating immigration should not occur, that is* $\hat{m} - \hat{a} = 0$.

Proof. We can substitute for a and m from the two consistency constraints into the planners problem (13). Instead of the inequality condition $m - a \geq 0$, we can alternatively use the conditions $C_n, C_a(\tau) \in [0, K]$ and $T_a \in [0, 1]$. Then the problem is to maximize a continuous function on a compact set, given the parameter values h, α, θ, L. For such problems the Theorem of the Maximum applies, and hence the optimal cutoff levels are upper hemi-continuous in the parameter values. Since by assumption the optimal cutoffs are unique, they are continuous in h, α, θ and L. And because a and m are continuous functions of the cutoffs, they are also continuous in the parameters.

[12] This is a much weaker requirement than having a unique solution to the first order conditions, because it allows for multiple *local* maxima.

It is easy to see from the first-order conditions that for h close enough to 1, there cannot be an interior solution. This is because in the interior case the cutoff C_n is strictly smaller than $1-h$ (see (14)). Thus there must be some high values of h where the optimum is corner. Finally, there is a unique value of h (not necessarily between zero and one) where $m = a$ solves the equations in (14). Substituting $a = m$ into $C_a(\tau)$ and C_n, solving $C_n = 0$ for m and using this value in (14) leaves us with the following equation:

$$\frac{2\alpha(1-\theta)}{1-h} - 1 = L\int_0^{1-h} F(1-h-\tau)\,dG(\tau). \tag{16}$$

The left-hand side of (16) is increasing in h, the r.h.s is decreasing. Also, at $h = 1$ the l.h.s is bigger than the r.h.s. Thus there is non-negative solution if the l.h.s is smaller than the r.h.s when $h = 0$, so $m = a$ solves (14) at some $0 < h < 1$ when:

$$\frac{2\alpha(1-\theta)}{L} < \int_0^1 F(1-\tau)\,dG(\tau). \tag{17}$$

Let us define $\hat{h}$ as the solution of (16) when it exists, otherwise let its value be zero. It is easy to see that in a corner solution $\hat{m}$ is decreasing in h, and that $\eta \geq 0$ iff $h > \hat{h}$. Thus the first-order condition for a corner optimum is satisfied only when $h \geq \hat{h}$. We only have to show that at these values of h the corner solution is indeed the optimum. But we already know this when h is close to $h = 1$, and we also know that $\hat{m}$ (and $\hat{a}$) are continuous in h. This means that as h decreases and the solution switches to an interior type, there must be an h where the two types coincide, and thus $m = a$ solves (14). But this can happen only at $h = \hat{h}$, which is also the lower limit of the possible corner optima. Thus for $h \geq \hat{h}$ we have a corner optimum, otherwise the optimum is interior. □

Thus when the planner's problem has a unique solution, it has the same structure as the equilibrium outcome. The only difference is that non-assimilating immigration might not be optimal even for very poor countries if they are small, cultural differences are large and non-assimilating immigrants are not very isolated in the North (see [17]). In fact it is easy to see that even in the case of an interior cutoff, we have $\hat{h} < \bar{h}$. This follows from the fact that the right-hand sides of (12) and (16) are the same, but the l.h.s. of the latter is bigger. Thus there is a range of source country income levels where there is non-assimilating immigration in equilibrium, but not in the optimal solution.

Comparing (9) and (15) reveals that the equilibrium outcome is globally optimal when $h \geq \bar{h}$. On the other hand, it is possible to show that an interior equilibrium cannot be optimal. Assume that it is, thus we have $\hat{m} = m^*$ and $\hat{a} = a^*$. Looking at the equations of $c_a(\tau)$ and $C_a(\tau)$ shows that in an interior case the latter is bigger than the former for all τ. Thus $\hat{a} = a^*$ only if $T_a < \tau_a$, from (8) and (14). But in this case we also have $C_n > c_n$, which together with $T_a < \tau_a$ implies that $\hat{m} - \hat{a} > m^* - h^*$, which leads to $\hat{m} > m^*$.

Thus an interior equilibrium is never optimal. We cannot say, in general, how the size and composition of equilibrium migration deviates from the optimal values. However, the comparison is relatively easy when $\hat{h} \leq h \leq \bar{h}$. In this case there is too much non-assimilating migration (there should be none) and not enough assimilating migration (because $\tau_a < 1 - h$). The following argument proves that overall migration is also too high:

$$
\begin{aligned}
m^* - \hat{m} &= F(c_n)\,[1 - G(\tau_a)] - \int_{\tau_a}^{1-h} F(1 - \tau - h)\,dG(\tau) \\
&= F(c_n) - \int_{\tau_a}^{1-h} f(1 - \tau - h)G(\tau)\,d\tau \\
&> F(c_n) - \int_{\tau_a}^{1-h} f(1 - \tau - h)\,d\tau \\
&= 0,
\end{aligned}
$$

where integration by parts leads to the second equality, and the inequality uses the fact that $G(\tau) < 1$.

By the continuity of the optimal and equilibrium solutions these inequalities also hold when h is not much smaller than $\bar{h}$. For lower values of h, however, much less can be said in general. It still "looks" likely that there is too much non-assimilating migration and too little assimilating migration, but we can only be sure if $L < 1$. In this case it is easy to check that $C_n < c_n$ and $T_a > \tau_a$, which together with $C_a(\tau) > c_a(\tau)$ guarantees that $\hat{a} > a^*$ and $\hat{m} - \hat{a} < m^* - a^*$. But overall migration can be either too high or too low even when $L < 1$.

Why do we have these ambiguities, given that we emphasized the negative externality caused by non-assimilating immigration on natives? The answer is that there is also a positive externality involved, which the planner takes into account when making her decision, and it arises in the migration decision. When an additional Southerner decides to migrate, she compares her costs to her benefits from living in the North. But when she moves, she also benefits non-assimilating immigrants, because they are more likely to get an efficient match. Thus in optimum the planner must balance the negative externality for natives with the positive externality for non-assimilating immigrants. This can lead to suboptimal migration levels for poor countries, and might even call for an increase in non-assimilating immigration when the South is poor *and* large.

To conclude this section, Proposition 4 summarizes the discussion on equilibrium and optimal migration. Equipped with these results, we can turn to the design of policies that can improve upon the equilibrium outcome.

Proposition 4. *For fairly developed source countries ($h \geq \bar{h}$) the level and composition of migration are socially optimal. Countries with an income range of $\hat{h} < h < \bar{h}$ send too many non-assimilating immigrants (instead of none) and not enough assimilating immigrants, and the overall migration level from such countries is too large. Finally, for poor countries it might be optimal to have non-assimilating immigrants, if the condition in* (17) *is satisfied. Even in this case, there are still too many non-assimilating immigrants and too few assimilating immigrants when $L < 1$, but any pattern might arise when $L > 1$.*

5 Policy evaluation

5.1 Implementing the first best

Having compared the optimal and equilibrium allocations, let us turn now to the implementation of the social optimum. The first important message of the model is that the optimal policy has to be conditioned on the model parameters. This is not just because the optimal and equilibrium solutions are functions of these parameters, but more importantly because the structure of the equilibrium and optimal outcomes changes at some parameter values. Because of these structural switches, it is best to talk about the different possibilities in turn.

The easiest case is when the source country is sufficiently rich[13] so that $h \geq \bar{h}(\alpha, \theta, L)$. In this case the equilibrium level and composition of migra-

[13] Although I do not always indicate, both $\bar{h}$ and $\hat{h}$ depend on the other parameters of

tion are globally optimal. There are no non- assimilating immigrants, and from (9) and (15) it is clear that the number of immigrants is the same in the two allocations. Thus for developed source countries there is no need for a policy intervention. Notice that this also holds from the point of view of the North, since natives there are indifferent to the size of assimilating migration. Therefore the equilibrium allocation is also (weakly) optimal for the North.

The next case is when the source country has a medium income level, $\hat{h} \leq h < \bar{h}$. It is worth noting that this range might extend to poor countries as well, because there is a set of parameter values with positive measure where $\hat{h} = 0$. In any case, for such source countries the policymaker needs to achieve two objectives: limit the size of overall migration and change its composition such that only people with low assimilation costs who will learn the culture of the North will migrate. Because of this latter goal, migration policy[14] has to target assimilating and non-assimilating immigrants differently. The optimal policy is a lump-sum tax on *non-assimilating* immigration, since from the planner's point of view their incentives to migrate are too high. The size of the tax D is given by the following equation:

$$D = 1 - h - \frac{\alpha(1-\theta)}{1+\hat{m}}, \tag{18}$$

where $\hat{m}$ is given by (15). The rational behind this policy can be seen from (4), since with the optimal tax D the cutoff for non-assimilating immigration is zero at the optimal migration level. When there is no non-assimilating immigration, $c_a(\tau) = 1 - h$, since all people who have positive gains from assimilating migration will move. Thus the selective tax both eliminates non-assimilating immigration and sets the migration level at its optimal value. As in the previous case, the outcome is also optimal from the Northern point of view, since non-assimilating immigration is zero. Finally, notice that the tax is actually never levied, and only serves as a deterrent. This means that we don't have to worry about allocating tax revenues, because there are none.

the model (α, θ and L). In the analysis, we keep parameters other than h constant when comparing a "poor" and a "rich" country. Thus we have to be careful about real world examples, because countries usually differ along more than one dimensions.

[14]When no ambiguity arises, I will use the term "migration policy" also to refer to assimilation policy.

Apart from some practical problems (such as how to verify assimilation status), there is an additional caveat in administering the tax that stems from the timing of the assimilation and migration decisions that is implicit in the model. To be more specific, the problem is that the policymaker might not credibly commit to the tax, since it is an out-of- equilibrium threat. If we think of the planner and a non-assimilating immigrant as playing a two-period game, the optimal outcome is a Nash-equilibrium of this game, but it is not sub-game perfect. This is because if the immigrant decides to come regardless of the tax, the planner has an incentive to revaluate her policy once the immigrant is in the country. At this point the composition of immigration is given, so the policymaker's decision is to select an optimal assimilation level. This second-stage problem is the same as in Lazear (1995), who shows that because of the externality on natives an assimilation subsidy is required, given that they are already in the country. Therefore the tax penalty threat is not credible, and non-assimilating immigration occurs regardless of it.

This conclusion changes if the planner is concerned about reputation, because of the game between her and immigrants is repeated over time. In this case it might be possible to commit to the optimal tax. Alternatively, legislation can tie the policymaker's hand when credible commitment is not possible. This line of argument closely resembles to the rules versus discretion debate in macroeconomics, so I will not pursue it further. But it is interesting to point out that the analysis above sheds new light on the debate about bilingual education for immigrant children in the United States. Opponents of such programs question not its aim to help the assimilation process of immigrant children, but its effectiveness to achieve this goal. In a book about US immigration (Borjas 1999b), George Borjas argues that bilingual education in fact makes assimilation more difficult, and hence should be abolished. But our argument points out that even if bilingual education helps children to assimilate by making their learning of English more gradual, it is an indication of the credibility problem faced by the policymaker. We can see that for each immigrant group already in the country there is an incentive for assimilation subsidies (such as bilingual education that is optimally administered), but for future immigration they should not be used. Thus the debate can be understood as a clash between the short-term and long-term interests of the country, without referring to the imperfections of the policy itself.

The final possibility is when the sending country is poor and the parameters are such that $0 < h < \hat{h}$ is a non-empty set. In this case some non-assimilating immigration is desirable, and the only thing we know in general is that the equilibrium allocation is not optimal. Let $\hat{m}$ and $\hat{a}$ be the optimal values for the overall number of immigrants and the size of assimilating immigration. Then the optimal transfers to the two immigrant groups, D_n and D_a, are given by[15]:

$$D_n = \frac{\alpha(1-\theta)(\hat{m}-2\hat{a}-1)}{(1+\hat{m})^2} \qquad \text{and} \qquad D_a = \frac{\alpha(1-\theta)(\hat{m}-\hat{a})}{(1+\hat{m})^2}. \tag{19}$$

It is easy to show that with these transfers the cutoff for assimilation is T_a, and the cutoffs in migration costs are the same as in the optimal allocation. The optimal policy calls for a subsidy for assimilating migration, because of the positive externality they generate for non-assimilating immigrants. Non-assimilating migration is more likely to be penalized, but if the optimal level of non-assimilating migration is large, they are also to receive a subsidy. In their case the planner balances the negative externality on Northern natives with the positive externality on fellow non- assimilating immigrants, so the sign ambiguity is natural.

Contrary to the previous case, there is a potential revenue (or deficit) for the planner, given by:

$$-[(\hat{m}-\hat{a})D_n + \hat{a}D_a] = \frac{\alpha(1-\theta)(\hat{m}-\hat{a})(1+\hat{a}-\hat{m})}{(1+\hat{m})^2}.$$

One possibility is to distribute the revenue (or the deficit) uniformly across all people, so that incentives are not distorted for the migration and assimilation decisions. Alternatively, the revenue (when it is positive) might be used to compensate Northern natives, whose own interest dictates no non-assimilating immigration. Since the welfare of Northern natives does not enter any decisions, subsidies given to them do not distort any incentives. Full compensation is possible if the revenue is greater than the loss natives suffer from non-assimilating immigration, which is given by $\alpha(1-\theta)(\hat{m}-\hat{a})/(1+\hat{m})$ (see the value function for natives). Unfortunately one can easily check that the tax revenue is always smaller than the native loss from immigration. Thus it is possible to alleviate, but not to fully compensate native loss.

The same caveats apply as in the case of middle income countries, at least when the transfer to non-assimilating immigrants is negative. Given that such an immigrant enters the North, the planner has an incentive to

[15]Note that here we define the transfers as subsidies, in contrast to the previous case.

deviate from the ex ante optimal policy. The problem is even more serious, because the North has to commit to a policy that is not in its best interest. When some compensation of Northern natives is possible, this can be partially overcome. Also, the two countries might cooperate (perhaps because of other shared goals) in immigration policy to approximate the social optimum. Otherwise, it is hard to see who enforces the first best, since policymaking is naturally a Northern affair.

The following proposition summarizes the results of this section about the optimal policies that achieve the first best. It is also interesting to look at some other popular policy instruments and evaluate their effect in the present framework. This is the topic of the next section.

Proposition 5. *The optimal policy for countries in the income range $\hat{h} \leq h < \bar{h}$ is a tax on non-assimilating immigration that is a sufficient deterrent for such migration. When some non-assimilating migration is desirable ($h < \hat{h}$), assimilating immigrants should receive a subsidy and non-assimilating immigrants a transfer that can be either negative or positive, depending on the model parameters. In both cases the problem of time inconsistency arises for the social planner. When there is non-assimilating migration in optimum, the likelihood of implementing the first best depends on the magnitude of compensation for Northern natives and on two-country cooperation.*

5.2 The effect of some other policies

There are various policies that have been advocated by the immigration literature. Now that we have identified the optimal policies in the present framework, it is instructive to see why some of these popular initiatives fail to implement the first best. To keep matters simple, I will concentrate on the case of middle income source countries ($\hat{h} \leq h < \bar{h}$), where the social optimum and the Northern interest coincide. In this case we can safely evaluate policies designed to maximize Northern welfare, which the literature focuses on.

The basic property of the optimal policy here is that it has to target the size and composition of the immigrant group at the same time. This means that any policy that focuses on only one of the two is doomed to fail. Let us discuss some possibilities in turn. First, countries might simply limit the amount of immigrants they are willing to receive. Even allowing for different quotas for different source countries (which is what optimality calls for), the composition of immigration won't be optimal. Assuming that those who benefit most will migrate (either because immigration permits are sold or because of some other form of rationing takes place), it is easy to see that at the optimal migration level some non- assimilating immigrants will come. This is because at this level of migration $c_n > 0$ whenever $h < \bar{h}$. Of course it is possible to set the migration limit so low that only assimilating migration occurs, but then their number will be suboptimal. Thus a policy that only targets the migration level cannot achieve the first best.[16]

The mirror image of setting the size of migration is influencing its composition. In his model of cultural interactions Lazear (1995) concludes that because assimilation is suboptimal, an assimilation subsidy is called for. In our framework of endogenous migration it is easy to see why such a policy fails to achieve its goal. Let us assume that instead of τ the immigrant needs to spend only $\tau - s$ time to assimilate. This leads to the following cutoffs in the equilibrium allocation:

$$\begin{aligned} \tau_a &= \frac{\alpha(1-\theta)}{1+m} + s \\ c_n &= 1 - h - \frac{\alpha(1-\theta)}{1+m} \\ c_a(\tau) &= 1 - h - \tau + s. \end{aligned}$$

We can immediately see that the first best cannot be implemented this way. The reason is that non-assimilating immigration is not penalized. Consequently, for any $h < \bar{h}$ we have $c_n > 0$ at the optimal migration level $\hat{m}$. In fact it is easy to show that for any positive subsidy the level of migration increases, so that c_n will increase as well. By subsidizing assimilation it is possible to push the assimilation cutoff to its optimal level of $1 - h$, but that will not eliminate non-assimilating immigration. Because of the positive effect of m on c_n, it is even possible that non-assimilating migration increases as a result of the subsidy. Thus (as we discussed earlier) such a policy might be optimal for an immigrant group already in the host country, but it does not achieve the first best when rational immigrants take into account the subsidy when they decide about moving.

[16] If we only care about Northern welfare this is not quite true, because the North is indifferent to the number of assimilating immigrants. But if immigration brings some gains for the North so that optimal immigration is strictly positive for the host country the same conflict between size and composition would arise.

So far we took the parameters as given and assumed that they cannot be chosen by the planner. It is interesting to relax this assumption and allow the planner to influence these values. An interpretation of this is that the parameters represent variables that are fixed in the short run, but can be changed in the long term. Thus we are interested in the design of a framework that allows for the best possible outcome in the migration problem. There is much scope for speculation here, and it is impossible to tackle all the issues. Thus I will only look at α, the level of segregation in the host country, and I will look at how the choice of this parameter influences the welfare of Northern natives. One can carry out similar exercises for the other parameters, but for lack of space I confine my attention here to α.

Some segmentation is inevitable, because people of different cultures choose to live separately. Still, there is some scope for policy to influence the amount of interaction between different groups. An example might be schooling. A government can encourage or discourage separation in schools, and there are examples to both in recent American history. Another example is Sweden, that has actively pursued a housing policy to spread immigrants evenly in the country (Murdie and Borgegård 1998). The question we ask here is whether more or less isolation is optimal for natives. Of course a full analysis should take into account immigrants as well, but as we will see the answer is not obvious even for natives and gives us some useful insights in itself.

Let us rewrite the native value function from Section 2.2. It is a function of the relative size of the non- assimilating immigrant group in the population, $(m-a)/(1+m)$. Using the notation ρ for this ratio, the Northern value function is:

$$V_N = 1 - \alpha(1-\theta)\rho.$$

It is not obvious that a tradeoff arises, because the direct effect of α is negative. There is, however, an indirect effect in equilibrium. In an optimal allocation this indirect effect could be ignored by the Envelope Theorem, but here the equilibrium is not optimal in general. Thus we have to see how a change in α influences the equilibrium value of ρ. In the Appendix I show that ρ is increasing in the level of segregation, so that native welfare depends on α in a non-trivial way.

The reason for this is easy to see. When the level of isolation increases, it is less likely that a native meets a non- assimilated immigrant, ceteris paribus. But more isolation also discourages assimilation, so that there are more non-assimilating immigrants in the Northern population. These two effect work in the opposite direction, so in general it is not obvious whether increasing or decreasing segregation is what natives should prefer. Thus it is very well possible that natives are better off in a less segregated society, even if we don't take into account ethical considerations.

6 Empirical evidence

In Section 3.2 we saw that the model has many unambiguous comparative statics predictions. In this section I will perform some preliminary tests that contrast these predictions with the data. I use the 1% sample of the 1990 US Census, which is large enough to have many immigrants and has information on their country of origin as well as their knowledge of English. The latter is very important since the most plausible proxy for assimilation is language proficiency. To try to confirm to the model's assumptions, the dataset only contains immigrants who arrived above the age of 18, for whom labor market characteristics can reasonably be thought of as predetermined. The data thus created contains 460,048 observations. Since not all the variables in the regressions have valid observations for all immigrants, the equations are estimated with around 330,000 observations.

Variable	Mean	Std. Dev.	Min	Max
AGE	46.00958	16.38753	20	90
AGECOME	32.18428	11.72949	19	89
YEARSCH	10.37372	5.024153	0	20
POP	.4786744	.989511	.0001665	4.472451
Y	.2548764	.166669	.01	1.65
DIST	6.86729	4.131057	.733894	16.37082

Table 1: Descriptive statistics for the continuous variables

The model's predictions refer to country level variables, such as the number of assimilating and non-assimilating immigrants, the source country's population, GDP and its cultural distance from the host country. Equivalently, one can interpret these predictions for individuals. For example, for a given person assimilation is more likely if she comes from a small country. The probability to assimilate (conditional on immigration) will be a function

SEX	%	ENGLISH	%	RHHLANG	%
Male	48.21	Not at all	14.86	Spanish	46.47
		Not well	25.03	Indo-Eur.	26.08
Female	51.79	Well	26.56	Asian	33.29
		Very well	33.63	Other	3.14

Table 2: Frequencies for the categorical variables

of the country level variables, and plausibly also of individual characteristics that proxy for assimilation costs and benefits.

Thus in order to test the model, I regress a measure of English knowledge on basic individual characteristics and information about the immigrant's source country. The former set of variables can be found in the Census, while the latter is matched from version 5.6 of the Penn World Tables (1994). The distance measure is the arc distance of the capital cities of the source countries from Washington, DC. Table 1 shows some descriptive statistics for the continuous variables: age, age of arrival at the US, years of schooling and the population (relative to the US), GDP (relative to the US) and distance from US (in 1000 km's) of the source country at the time of arrival. Table 2 provides the frequency distributions for the categorical variables: gender, English knowledge and the main household language. Almost half of the immigrants come from Spanish-speaking countries, and around 40% of them speaks English not well or not at all.

There are four regressions reported in Table 3. The dependent variable in the first two is the categorical variable ENGLISH , and in the second two the binary variable ASSIM. The latter is calculated from ENGLISH by defining an immigrant assimilated if he or she speaks English well or very well. For both dependent variables regressions are reported with or without household language dummies. Together with DIST they proxy for the cultural distance between the US and the source country. The difference between the former and the latter is that DIST must also correlate with the cost of moving, whereas the language dummies might capture learning costs as well. Thus the theoretical predictions are more clear with DIST.

In all four regressions the predictions of the model are born out and a very few variables explain a large portion of the variation in the data. An immigrant is more likely to assimilate if he or she comes from a larger, less advanced and geographically closer country. For example, a 1% increase

Dependent variable	ENGLISH	ENGLISH	ASSIM	ASSIM
CONSTANT	0.547	1.114	0.0712	0.341
	(81.1)	(84.0)	(22.3)	(49.2)
POP	-0.080	-0.082	-0.039	-0.040
	(-48.8)	(-49.6)	(-49.9)	(-50.6)
Y	0.309	0.087	0.122	0.028
	(32.1)	(8.7)	(26.4)	(5.7)
DIST	0.060	0.054	0.027	0.024
	(137.9)	(66.9)	(130.6)	(61.9)
SEX	-0.102	-0.098	-0.044	-0.042
	(-34.2)	(-33.3)	(-30.8)	(-29.9)
AGE	0.026	0.022	0.011	0.009
	(195.6)	(161.9)	(172.6)	(142.6)
AGECOME	-0.038	-0.036	-0.016	-0.015
	(-206.5)	(-195.7)	(-181.0)	(-171.0)
YEARSCH	0.085	0.081	0.035	0.033
	(255.1)	(240.0)	(220.4)	(206.4)
SPANISH		-0.427		-0.181
		(-44.8)		(-39.4)
INDOEUR		-0.083		-0.034
		(-9.4)		(-7.9)
ASIAN		-0.411		-0.166
		(-44.5)		(-37.3)
R^2	0.365	0.384	0.310	0.326

Table 3: Individual assimilation regressions

in the relative population of the source country leads to a 4% decrease in the probability to assimilate. A remarkable fact is the robustness of the distance parameter to the inclusion of the language family dummies. In all specifications DIST is highly significant and has the correct sign, and the parameter value is meaningful. It is obviously important to incorporate the language family variables, as the second and fourth specifications show. Ceteris paribus, immigrants from Asia or Latin America are less likely to assimilate, and the effects are very large. Interestingly, the Asian and Spanish dummies are very similar in size. This shows that keeping everything else constant, Asian and Latino immigrants have the same propensities to assim-

ASSIM	Spanish	Indo-European	Asian	Other	Total
0	58.41	20.98	29.06	12.75	40.08
1	41.59	79.02	70.94	87.25	59.92

Table 4: Assimilation by language groups

ilate. On the other hand, Table 4 reveals that the unconditional distribution of the assimilation measure is very different for these two language groups - Asians are much more likely to learn English. The regressions show that there is absolutely no mystery about this finding, and we can explain the difference completely by observable individual and country characteristics.

It must be kept in mind that the empirical results presented are quite preliminary. One particularly interesting direction is to try to estimate the model with the structural break. That might explain the large negative coefficients for Asians and Spanish speakers, since Indo-Europeans are likely to come from high-income countries and thus their assimilation patterns are different from the other two groups. This might hold for the "Other" category as well, since their numbers are fairly small. This and other extensions of the simple regressions look very promising, but out of the scope of the present paper. Nevertheless, the results reported above strongly support the basic findings of the paper, and call for a separate and more detailed empirical project along the model's lines.

7 Conclusion

Cultural differences are an important aspect of the immigration process, not least because natives seem to care a great deal about them. This paper has developed a model that explores the role of culture in the migration and assimilation decisions. The model enabled us to look at the welfare properties of the equilibrium level and composition of immigration. Given the external effects of immigrant decisions on natives, it is not surprising that in general the equilibrium outcome is inefficient. But this is not always the case, and in fact we saw that there is set of parameter values (with positive measure) where the optimal and equilibrium allocations coincide.

Given the structural switches in the optimal and equilibrium outcomes, the policy implications differ for different source countries. In most cases when the equilibrium is inefficient a tax on non-assimilating immigrants implements the first best, although it might be hard for the policymaker to commit to such a policy. Also, for some parameter values natives in the host country and the social planner (who cares about global welfare) have different objectives. When the revenue from the optimal tax can be used to compensate natives, this problem can be alleviated, if not completely overcome. Otherwise, since policy decisions are made in the host country, implementing the social optimum might not be possible in practice.

To make the model tractable I abstracted away from many issues that traditionally interest researchers of migration. Apart from assimilation skills, people do not differ in their labor market abilities. There is only one (abstract) sector of production, and there is no role for capital (physical or human) in the model. Also, the analysis is essentially static, although there is an implicit dynamic element: the timing of the migration and the assimilation decisions. Still, we could make the model more explicitly dynamic by allowing for subsequent immigrant generations, and look at the evolution of their composition over time. Clearly, incorporating these would lead to additional insights. But they would also make the model even more complicated, and I believe the main conclusions would not change substantially. Thus in my view the present model is an acceptable compromise between tractability and usefulness.

Even so, there are several issues that would be interesting to look at in subsequent work. One of them concerns the benefits immigrants bring to the host country. An interesting issue is the possibility that immigrants create trade with their source country. Casella and Rauch (1998) develop a model where ethnic networks facilitate trade between countries by having access to information not available through the market. Ethic groups, such as immigrants, can thus alleviate the market failure caused by incomplete information. There is also some empirical evidence that immigrants act as catalysts for foreign trade. Helliwell (1997) documents that trade between US states and Canadian provinces is larger if they have a larger share of each other's citizens as immigrants. In another work, Head, Ries and Wagner (1997) examine the effect of immigration on trade between Canada and various East Asian countries, and they find a positive relationship.

The crucial question with this and other possible benefits is whether non-assimilating immigration generates positive externalities for natives. If not, we can safely abstract away from them, since the qualitative nature of the problem remains the same, though the optimal levels might change. In the trade creation context, this would require immigrants to appropriate all the benefits from improved efficiency, or at least that they generate positive externality only for their source country. I think this is not unreasonable to assume, and it is possible to extend the present model to generate such externalities. Since this does not change the qualitative conclusions of the paper much, I leave the extension for future research.

Another important issue is that the assimilation decision might involve strategic considerations if natives can also learn the culture of immigrants. In this case learning by one group decreases the incentives for learning for the other. In the present paper I avoided the strategic aspect of learning by not allowing natives to "reverse assimilate". First, it would make the analysis too complicated for analytical results. Second, I believe that in the case of modest immigrant flows this is a fairly innocuous simplification. Culture in a country includes not just language, but habits, norms, and the legal and political systems. Unless the immigrant group is large enough to form its own self-governing society (in which case the simple random matching framework is irrelevant anyway), a large chunk of culture has to be acquired by immigrants regardless of native decisions. Furthermore, the strategic view is important if the two groups are similar in size. Thus if immigration is modest, not allowing "reverse assimilation" can be viewed as a good approximation to reality. In other cases this need not be the case, for example when we look at the role of culture in inter-country relations. To address this issue, in a recent paper (Kónya 2000) I incorporate strategic learning decisions into the analysis in a model of international trade and culture.

Finally, further empirical investigations into the role of culture in the migration process are desirable. The empirical findings reported in the paper are very encouraging, but more needs to be done. A very important (and very difficult) task would be to quantify the extent of the cultural externalities between natives and immigrants. The qualitative conclusions of the present model do not depend on this, but we need to understand the practical importance of the problem better. Historical evidence – such as the examples cited in Sowell (1996) – is available, but very hard to measure. It is not clear how we can provide more quantitative measures, but it would clearly be very important.

There are many ways to extend the model, and the possibilities I mentioned above are some of them. Empirical research is clearly a priority, if only for the recent interest of politics in immigration and culture, especially in Europe. Although there is scope for policy interventions, there is a clear danger that in a heated political atmosphere they would be overdone. Migration of poor people into rich countries is desirable for immigrants (and possibly – in a more realistic setting – for the host country), and that makes the implementation of right policies even more important. Some policies that seem attractive might not have the desired effect, and they can even worsen the problem instead of solving it. This paper hopefully provides a first step in the direction of avoiding such costly mistakes.

References

Altonji, J. G. and Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives, *in* J. M. Abowd and R. B. Freeman (eds), *Immigration, Trade and the Labor Market*, University of Chicago Press.

Auerbach, A. J. and Oreopoulos, P. (1999). Generational accounting and immigration in the United States, *Working Paper 7041*, NBER.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants, *American Economic Review* **77**(4): 531–552.

Borjas, G. J. (1999a). The economic analysis of immigration, *in* O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol. 3A, North-Holland.

Borjas, G. J. (1999b). *Heaven's Door*, Princeton University Press.

Burda, M. and Wyplosz, C. (1992). Human capital, investment and migration in an integrated Europe, *European Economic Review* **36**: 677–684.

Card, D. and DiNardo, J. E. (2000). Do immigrant inflows lead to native outflows?, *Working Paper 7578*, NBER.

Casella, A. and Rauch, J. E. (1998). Overcoming informational barriers to international resource allocation: prices and group ties, *Working Paper 6628*, NBER.

Chiswick, B. R. (1978). The effect of Americanization on the earnings of foreign-born men, *Journal of Political Economy* **86**: 897–921.

Hanson, G. H. and Slaughter, M. J. (1999). The Rybczynski theorem, factor-price equalization and immigration: evidence from U.S. states, *Working Paper 7074*, NBER.

Head, K., Ries, J. and Wagner, D. (1997). Immigrants as trade catalysts, *in* A. Safarian and W. Dobson (eds), *The people link: Human resource linkages across the Pacific*, University of Toronto Press.

Helliwell, J. F. (1997). National borders, trade and migration, *Working Paper 6027*, NBER.

Kónya, I. (2000). Modeling cultural barriers in international trade, Mimeo, Northwestern University.

Layard, R., Blanchard, O., Dornbusch, R. and Krugman, P. (1992). *East-West Migration*, The MIT Press, Cambridge, MA.

Lazear, E. P. (1995). Culture and language, *Working Paper 5249*, NBER.

Lazear, E. P. (1996). Culture wars in America, *Essays in public policy*, Hoover Institution, Stanford University.

Markusen, J. R. and Zahniser, S. (1997). Liberalization and incentives for labour migration: theory with applications to NAFTA, *Working Paper 6232*, NBER.

Murdie, R. A. and Borgegård, L.-E. (1998). Immigration, spatial segregation and housing segmentation of immigrants in metropolitan Stockholm, 1960-95, *Urban Studies* **35**(10): 1869–1888.

Sowell, T. (1996). *Migrations and cultures: a world view*, BasicBooks.

Storesletten, K. (2000). Sustaining fiscal policy through immigration, *Journal of Political Economy* **108**(2): 300–323.

Trefler, D. (1997). Immigrants and natives in general equilibrium trade models, *Working Paper 6209*, NBER.

Venables, A. J. (1999). Trade liberalisation and factor mobility: an overview, *in* R. Faini, J. D. Melo and K. F. Zimmermann (eds), *Migration: the controversies and the evidence*, Cambridge University Press.

Zhou, M. and Bankston III, C. L. (1998). *Growing up American: How Vietnamese Children Adapt to Life in the United States*, Russell Sage Foundation, New York.

Appendix

The comparative statics analysis is straightforward in the case of a corner solution, since the left-hand side of (9) does not depend on m. Thus the migration level in this case is increasing in L and it is decreasing in h.

For an interior equilibrium, let us define A and B with the following equations:

$$A = \int_0^{\tau_a} F(1-h-\tau)\, dG(\tau)$$
$$B = F(c_n)[1-G(\tau_a)],$$

where

$$\tau_a = \frac{\alpha(1-\theta)}{1+m}$$
$$c_n = 1-h-\frac{\alpha(1-\theta)}{1+m}.$$

Thus we have that $m = A + B$ and $m - a = B$, and it is easy to see that $A_m < 0$ and $B_m > 0$. Assumption 1 also guarantees that $1 - A_m - B_m > 0$.

The effect of parameters on m is given by the expression $(A_j + B_j)/(1 - A_m - B_m)$, where $j = \alpha, \theta, h, L$. Since the denominator is positive, the sign of the expression depends on the sign of the numerator. It is immediate that $A_h, B_h < 0$ and $A_L, B_L > 0$, hence m is increasing in L and it is decreasing in h. For α we have that:

$$\frac{\partial(A+B)}{\partial\alpha} = f(c_n)[1-G(\tau_a)]\frac{\partial c_n}{\partial\alpha} < 0,$$

so m is increasing in the level of segregation. Similarly for θ :

$$\frac{\partial(A+B)}{\partial\theta} = f(c_n)[1-G(\tau_a)]\frac{\partial c_n}{\partial\theta} > 0,$$

so m is increasing in the level of cultural similarity.

For the other results we can use the following matrix equation:

$$\begin{bmatrix} 1 - A_m - B_m & 0 \\ -A_m & 1 \end{bmatrix} \begin{bmatrix} d\,m \\ d\,a \end{bmatrix} = \begin{bmatrix} A_j + B_j \\ A_j \end{bmatrix} d\,j,$$

where again $j = \alpha, \theta, L, h$. Then the derivatives of $m - a$ are given by:

$$\frac{\partial(m-a)}{\partial j} = \frac{B_j(1 - A_m) + A_j B_m}{1 - A_m - B_m}.$$

This expression can be easily signed when A_j and B_j have the same sign, as in the case of L and h. For α it is easier to evaluate the derivative of a:

$$\frac{\partial a}{\partial \alpha} = \frac{A_\alpha(1 - A_m - B_m) + A_m(A_\alpha + B_\alpha)}{1 - A_m - B_m} > 0.$$

Since the level of migration is decreasing in α and assimilating migration is increasing in it, it is necessary that $m - a$ decreases with α. Finally, in the policy section I claimed that $\rho = (m-a)/(1+m)$ is decreasing in α. Now we can see why this is the case, given that we can rewrite ρ as $1-(1+a)/(1+m)$.