Intergenerational Assimilation of Minorities: The Role of the Majority Group

By

Ryo Itoh (Tohoku University)
Yasuhiro Sato (The University of Tokyo)
Yves Zenou (Monash University)

December 2021

CREPE DISCUSSION PAPER NO. 114

CENTER FOR RESEARCH AND EDUCATION FOR POLICY EVALUATION (CREPE)
THE UNIVERSITY OF TOKYO
http://www.crepe.e.u-tokyo.ac.jp/

Intergenerational Assimilation of Minorities: The Role of the Majority Group*

Ryo Itoh[†] Yasuhiro Sato[‡] Yves Zenou[§]

December 20, 2021

Abstract

We develop a dynamic model of assimilation of ethnic minorities that posits a tradeoff between higher productivity and wages and greater social distance to the culture of origin. We also highlight the importance of the assimilation of the past generation and the role of the majority group in the assimilation of ethnic minorities. First, there is an inverted $U-$shaped relationship between the degree of tolerance of the majority individuals and the average level of assimilation in the society. Second, more tolerance from the majority group generates positive externalities for the minority group, while each minority's individual assimilation effort affects the welfare of the majority individuals differently depending on the initial minority assimilation level. Finally, the more the majority individuals are tolerant toward the minority group, the more the minority individuals will assimilate to the majority group, while the reverse is not always true. In fact, when there is too much assimilation, the majority group may reduce its degree of tolerance toward the majority group.

Keywords: Identity, assimilation, social norms, group status, dynamics, welfare, majority's acceptance.

JEL Classification: J15, Z13.

*This work was supported by JSPS KAKENHI, Grant Number 18H00842.

[†]Graduate School of Information Sciences, Tohoku University, Japan. Email: itoh@se.is.tohoku.ac.jp.

[‡]Faculty of Economics, University of Tokyo, Japan. Email: ysato@e.u-tokyo.ac.jp.

[§]Department of Economics, Monash University, Australia. Email: yves.zenou@monash.edu.

1 Introduction

The United States and the European Union (EU) have a long history of accepting immigrants. According to the Pew Research Center,[1] the United States has more immigrants than any other country in the world. Today, more than 40 million people living in the U.S. were born in another country, accounting for about one-fifth of the world's migrants. In 2018, the U.S. foreign-born population reached a record 44.8 million. According to a recent report from the European Commission,[2] on January 1, 2020, 447.3 million inhabitants were living in the EU. Among them, 23 million were non-EU citizens (5.1% of the EU's total population), and nearly 37 million people were born outside the EU (8.3% of all EU inhabitants).[3]

A central concern is the integration and assimilation of these immigrants into the host country, not only of the first generation, but also of subsequent generations (Hammarstedt, 2009; Algan et al., 2010; Drouhot and Nee, 2019; Maskileyson et al., 2021). In particular, an important aspect of the immigrants' assimilation that has been neglected is the possible unwillingness of the majority group to accept immigrants in the host country.

Accordingly, the aim of this paper is to develop a dynamic intergenerational model of immigrants' assimilation that highlights not only the importance of the assimilation of the previous generation of immigrants but also the role of the majority group in promoting or hindering the assimilation of the current generation of immigrants. There are empirical papers showing that natives have an unfriendly stance toward immigrants and that immigration flows fuel these attitudes (Dustmann et al., 2019; Hangartner et al., 2019; Moriconi et al., 2019; Tabellini, 2020; Steinmayr, 2021). However, to the best of our knowledge, this is the first paper that investigates (at least in economics) how natives' (friendly or unfriendly) attitudes toward immigrants affect their assimilation and vice versa.

To understand the different forces at work, we first develop a model in which the majority's tolerance decision toward the minority group is not incorporated. We mainly focus on the intergenerational assimilation decision of minorities. The decision of assimilation to the majority's norm is decided by trading off higher productivity and wages with a greater social distance to their culture of origin and group status. Individuals are heterogenous with respect to α_i, the utility weight each minority places on incomes. In the static model, we show that a higher average assimilation effort of the previous generation results in a higher current average income and a higher assimilation effort only for individuals with a high α_i. We also show that the assimilation effort chosen by minority in-

[1] https://www.pewresearch.org/fact-tank/2020/08/20/key-findings-about-u-s-immigrants/.

[2] https://ec.europa.eu/info/strategy/priorities-2019-2024/promoting-our-european-way-life/statistics-migration-europe_en.

[3] For an overview on the trends, policies, and empirical evidence on immigration in Europe, see de la Rica et al. (2015).

dividuals is inefficiently high. Indeed, when a minority individual decides her assimilation effort, she ignores the effects of this choice on the ethnic minority's average income, which, in turn, negatively affects the utility of the other minority individuals by lowering the incentive to assimilate further. Thus, the effects on her own and other individuals' utility work as negative externalities and make the assimilation effort inefficiently high.

We then consider a dynamic model in which the assimilation effort of the current-generation minority is affected by the average assimilation effort of the previous generation. In the symmetric case where all individuals have the same α_i, we are able to characterize all stable steady-state equilibria. When the utility weight on the perceived distance is small enough, we show that there exists a unique equilibrium that can be either interior or corner (all minorities totally assimilate to the majority's norm). On the contrary, when this is not the case, multiple equilibria prevail because the perceived distance effect becomes more important compared to the role of the group status and direct cost. In this dynamic model, the minority's assimilation effort can be inefficiently high or low because each individual effort exerts negative externalities on the other individuals of the same generation, since more assimilation makes minorities care more about the decrease of their relative income status compared to that of the majority group, but positive externalities on the next generation, since more assimilation effort decreases the assimilation cost of the next generation. We then consider the asymmetric case when α_i is different for the minority individuals. In this case, we show that it is less likely that multiple equilibria will prevail.

In the second part of the paper, we explicitly model the majority's tolerance behavior toward the minority group. There is a tradeoff for each individual of the majority group. By increasing her degree of tolerance, the majority individual reduces the disutility of the perceived distance between their level of assimilation (the highest one) and that of the minority group. However, there is a direct cost of the tolerance effort that reduces their utility. What is interesting here is the feedback loop between the majority tolerance effort and the minority's assimilation effort. Indeed, the minority's average assimilation effort affects the majority's tolerance decision toward minorities, which, in turn, affects the individual decision to assimilate (which depends on the majority's average acceptance) and thus also has an impact on the minority's average assimilation effort.

In the static model, we show that there is an inverted $U-$shaped relationship between the degree of tolerance of the majority individuals and the average level of assimilation in the society. Indeed, when there is very little assimilation, there is no communication between groups. Therefore, the majority group finds it profitable to increase their tolerance toward the minority group. On the contrary, when there is too much assimilation in the society, it becomes too costly to increase tolerance, given the resulting low benefits.

In the dynamic model, we first show that the introduction of a majority acceptance decision weakens the convexity of the minority assimilation, reducing the possibility of multiple equilibria. Then, we study the efficiency of these equilibria. First, more tolerance from the majority group generates positive externalities on the minority group because it lowers their assimilation costs. In contrast, each minority individual assimilation effort affects the welfare of the majority individuals differently depending on the initial minority assimilation level. Interestingly, the efficiency of the equilibria depend on the ratio of the number of majority individuals relative to minority individuals, since it determines the strength of these externalities on majority individuals.

Finally, for the policy implications of the model, we show that when the minority group is well-assimilated, they generate positive externalities for the majority group, which increases the majority's welfare. Thus, the planner wants to increase assimilation in the country. However, this does not necessarily benefit the current generation's minority group because it reduces their relative status in the society but increases the welfare of the next generation by reducing their assimilation costs.

1.1 Related literature

There is an important *empirical* literature showing different significant influences on the assimilation process for immigrants: the quality of immigrant cohorts (Borjas, 1985), country of origin (Borjas, 1987, 1992; Chiswick and Miller, 2011), ethnic concentration (Edin et al., 2003; Damm, 2009), personal English skill (Chiswick and Miller, 1995, 1996; Lazear, 1999; Dustmann and Fabbri, 2003), and social networks (Xu, 2017; Biavaschi et al., 2021; Boucher et al., 2021).

There is also a growing *theoretical* literature that shows that the assimilation of immigrants has been impacted by their cultural identity (Akerlof and Kranton, 2010; Benabou and Tirole, 2011; Bisin et al., 2011; Panebianco, 2014; Bisin et al., 2016; Carvalho, 2016; Kim and Loury, 2019; Munoz-Herrera, 2021), their geographical location and their community cohesion (Bezin and Moizeau, 2017; Stark et al., 2018; Sato and Zenou, 2020), their cultural leaders (Hauk and Mueller, 2015; Prummer and Siedlarek, 2017; Verdier and Zenou, 2018), and their social networks (Verdier and Zenou, 2017; Boucher et al., 2021).

Our contribution to this literature is to develop a dynamic model of assimilation of ethnic minorities that highlights the tension between the benefits of assimilation, that is, higher productivity and wages, and its costs, that is, greater social distance to their culture of origin and group status, the importance of the assimilation of the past generation, and, more importantly, the role of the majority group in the assimilation of ethnic minorities; particularly, the two-way interaction between the minority's assimilation decision and the majority's tolerance of the minority group. To the best of our knowledge, at least from a theoretical perspective, this is the first paper that addresses these issues in a unified theoretical framework.

The rest of the paper unfolds as follows. In the next section, we provide some simple evidence (correlations) showing the determinants and dynamics of assimilation and the role of the majority group in the assimilation of immigrants. In Section 3, we focus on the dynamics of assimilation without modeling the role of the majority group. In Section 3.1, to understand the different forces at work, we develop a static model, and then, in Section 3.2, we analyze the dynamic model in which the previous generation of immigrants affects the current one. In Section 4, we introduce the role of the majority group's acceptance of the minority group in the assimilation process of immigrants and examine its policy implications. Finally, Section 5 concludes this work. In Appendix A, we provide a detailed analysis of the stylized facts displayed in Section 2. In Appendix B, we provide all the proofs of the results in the main text.

2 Stylized facts

The aim of this section is to provide some simple evidence on the dynamics of assimilation and the role of the majority group in the minority's assimilation, which, as we have seen in the Introduction, is not well documented. Appendix A develops in detail these analyses.[4] Here, we summarize the main relevant results.

Let us start with Section A.1, where we examine the factors that affect the assimilation of immigrants. Using the 2019 American Community Survey (ACS), we measure assimilation by the degree of the English fluency of immigrants.

First, we want to understand the *dynamics* of assimilation by investigating the relationship between the current assimilation (English fluency in 2019) of an immigrant and the share of assimilated immigrants in the past (English fluency in 2010) who live in the same county. Tables A2 and A3 show a strong, significant, and *positive* relationship between these two variables. This means that there is a *persistence* of assimilation over time (previous generations affect new generations) and thus motivates a *dynamic analysis of assimilation* in the theoretical section.[5]

Second, we investigate how the *cost* of assimilation affects the assimilation of immigrants. We measure this cost by the linguistic distance between the immigrant language and English. The higher this distance, the more costly it is to assimilate. In tables A2 and A3, we show that there is a strong, significant, and *negative* relationship between individual assimilation (English fluency in 2019) and linguistic distance.

In Section A.2, we investigate what affects the majority's acceptance of the minority culture.

[4]Observe that in this section, we never claim any causality, just correlations.

[5]In Section A.3, using another dataset (the International Social Survey Program) for 31 countries in 1995, 2003, and 2013, we show in Section A.3.2 that past assimilation is also strongly positively correlated with current assimilation (see Table A7).

Using the 2019 American Community Survey (ACS), we measure this acceptance by the county *intermarriage* rate between non-Hispanic whites and any other race. In tables A4 and A5, we show that there is a significant and inverted $U-$shaped relationship between the assimilation of immigrants (measured by either their English fluency or their naturalization rate) and the majority's degree of acceptance of the minority culture (measured by their intermarriage rate). In other words, an increase in the minority's assimilation rate raises the majority's degree of acceptance when assimilation is low while it lowers it when assimilation is high. Indeed, in these tables as well as in figures A1 (English proficiency) and A2 (naturalization), we see that the marginal effects change from positive to negative values as the immigrant's assimilation becomes larger.[6] This confirms one of our main theoretical results in Section 4: there is an inverted $U-$shaped relationship between the degree of tolerance of the majority individuals and the average level of assimilation in the society.

We would now like to develop a model that has these features: costs and dynamics of assimilation and the importance of the majority's degree of acceptance in the assimilation process.

3 Minority's assimilation decision without the majority's influence

To understand the assimilation decision of minority individuals, we first develop a baseline model in which we do not model how the majority group affects this assimilation effort. We start with a *static* model (Section 3.1) and then analyze the *dynamic* assimilation process (Section 3.2).

3.1 The static model

In this section, we fix the assimilation level of the previous generation and study how assimilation decisions can be transmitted to the next generation.

3.1.1 Social groups

Consider an economy with two groups of individuals: the *minority* and the *majority* groups. The number of minority individuals is m whereas the number of majority individuals is not explicitly given but is assumed to be larger than m. If we think of ethnicity, then the minority group is the *ethnic minority* group or the immigrant group while the majority group corresponds to the *native majority* group.

[6] In Section A.3, using another dataset (the International Social Survey Program) for 31 countries in 1995, 2003, and 2013, we show in Section A.3.3 that there is also an inverted $U-$shaped relationship between the assimilation of the immigrants (measured by the answer to the question: "How important is it for you (minority individual) to be able to speak the main language of the country where you live?") and the majority's degree of acceptance (measured by the answer to the question: "How important is it for you (majority individual) that minorities preserve their traditions?"). See Table A8 and Figure A3.

These two social groups, minority and majority, are "categories" that individuals learn to recognize while growing up (Akerlof and Kranton, 2010). Each individual is exclusively a member of one of these two groups. These groups are given, and we focus on the effort decision to assimilate for the ethnic minority group, that is, the degree to which they want to assimilate to the majority group. Quite naturally, we assume that the majority group is "totally"' assimilated, so that their "effort" to assimilate is exogenously given by $\overline{x}$. Each minority individual $i = 1, 2, \cdots, m$ can choose an assimilation effort $x_i \in [0, \overline{x}]$ by paying an effort cost. The effort x_i could be, for example, interpreted as the proficiency of the majority's language, so that $\overline{x}$ corresponds to being fluent in this language. Many papers have shown the importance of language in the assimilation and integration of immigrants (Chiswick and Miller, 1995; Lazear, 1999; Arendt et al., 2020; Boucher et al., 2021).

3.1.2 Production and wages

In the economy, the numéraire good is produced by only using labor. The production of this good exhibits constant returns to scale. The output per capita of the majority individuals is exogenously given by Z[7] whereas the output per capita of a minority individual i depends on her assimilation effort $x_i \in [0, \overline{x}]$ and is given by $y(x_i)$. We assume that $y(x_i)$ is twice continuously differentiable, increasing (i.e., $y'(x_i) > 0$), weakly concave (i.e., $y''(x_i) \leq 0$), and $Z \geq y(\overline{x})$. Indeed, $y'(x_i) > 0$ means that the more assimilated an ethnic minority individual (for example, the better she speaks the native language), the more she has access to better job opportunities and the higher her productivity and wages. This is well documented empirically. For example, Meng and Gregory (2005) have shown that intermarried immigrants, who are more assimilated, earn significantly higher incomes than endogenously married immigrants, even after human capital endowments and the endogeneity of intermarriage are taken into account. Using a different definition of assimilation, in the United States, Biavaschi et al. (2017) find that low-skilled migrants who Americanized their names experienced larger occupational upgrading than those who did not. Similarly, Arai and Skogman Thoursie (2009) examine data on immigrants who changed their surnames to Swedish-sounding or neutral names during the 1990s in Sweden. They find that there is a substantial increase in annual earnings after a name change. Finally, using language skills in Canada, Li (2013) shows that assimilated minority workers earn as much as the majority workers and much more than non-assimilated minority workers.

3.1.3 Social identity

We assume that three main factors affect the social identity and thus the socialization process in terms of the assimilation of ethnic minority individuals (Akerlof, 1997; Shayo, 2009, 2020; Sambanis and Shayo, 2013). First, each individual is aware of the different social groups or categories (i.e.,

[7]Since all majority individuals have the same output per capita, Z is also the *average* output per capita.

majority and minority groups) that exist in the society and know Z and $y(x_i)$. Second, each minority individual is a conformist and, thus, wants to minimize the *perceived distance* between x_i and that of the majority group. Third, each individual cares about the *status* of each social group, such that higher status implies higher utility.

Perceived distance

The concept of perceived distance and its adoption to the process of identification originated in the literature of categorization in cognitive psychology (Nosofsky, 1986; Turner et al., 2010). It has also been modeled by economists where agents are conformists and pay a higher cost when the perceived distance between their action and that of their peers (social norm) increases (Akerlof, 1997; Shayo, 2009; Patacchini and Zenou, 2012; Sambanis and Shayo, 2013; Liu et al., 2014; Boucher, 2016; Stark et al., 2018; Sato and Zenou, 2020; Ushchev and Zenou, 2020).

The *social norm* or reference point for each group is determined by $x = \overline{x}$ for the majority group and $x = 0$ for the minority group. Each minority individual i chooses the assimilation effort x_i, and this affects her "empathy" or taste for conformity $p(x_i)$ for the the majority's social norm. In other words, the perceived distance of ethnic minority i is given by:

$$p(x_i)(\overline{x} - x_i). \tag{1}$$

where $p(x_i)$ is a continuously differentiable and increasing (i.e., $p'(x_i) > 0$) function, with $p(0) = 0$ and $p(\overline{x}) = 1$.

Equation (1) thus assumes that, if an ethnic minority makes zero effort (i.e., $x = 0$) to assimilate to the majority's norm and thus lives in accordance to her own culture, her perceived distance is the lowest and equal to zero since $p(0) = 0$. Similarly, if an ethnic minority i chooses to fully assimilate to the majority's norm, that is, $x_i = \overline{x}$, her perceived distance is again the lowest and equal to zero.

Ethnic minority individuals are born with their social attitude bequeathed by their parents, that is, the average assimilation degree of the previous generation, X_0, which also determines their empathy or conformity to the majority's culture.

Group status

The last part of the utility function of an ethnic minority i includes a component related to the *status* of the group. The status of the group is determined by the difference between Z, the (exogenous) average income of the majority group, and $Y = \sum_{i=1}^{m} y(x_i)/m$, the (endogenous) average income of the ethnic minority individuals, weighted by $p(x_i)$, that is,

$$p(x_i)(Z - Y).$$

When choosing her assimilation effort x_i, quite naturally, we assume that each individual minority takes Y as given. Thus, since $p'(x_i) > 0$, the higher is an individual's assimilation effort x_i, the more she cares about $Z - Y$, the status of her group relative to that of the majority group. Observe that an increase in Y raises the social status of the minority group and thus decreases $Z - Y$; this lowers the incentive to assimilate. In other words, the higher the status of the minority group in the society, the lower the marginal utility of assimilation.[8] [9]

3.1.4 Utility function

Let us put the three parts of the utility function together. The utility function of a minority individual i choosing x_i is then equal to:

$$U(\alpha_i, x_i) = \alpha_i \underbrace{y(x_i)}_{\text{income}} - \rho \underbrace{p(x_i)(\overline{x} - x_i)}_{\text{perceived distance}} + \sigma \underbrace{p(x_i)(Z - Y)}_{\text{group status}} - \underbrace{h(X_0)c(x_i)}_{\text{effort costs}}. \tag{2}$$

where $\rho > 0$, $\sigma > 0$, and X_0 is the average assimilation effort of the previous generation, which is assumed to be fixed in this section, but will be endogenized in the next section. The first term of (2) represents the utility from own income (the material return), which is increasing in assimilation effort x_i. The second term captures the disutility from deviating from the social norm of the majority group (the perceived distance) while the third term is the payoff from the group status, which is also increasing in own effort x_i. The last term, $h(X_0)c(x_i)$, corresponds to the costs of assimilation effort x_i. Quite naturally, we assume that $c(x_i)$ is twice continuously differentiable, increasing ($c'(x_i) > 0$), and strongly convex ($c''(x_i) > 0$) while $h(X_0)$ is continuously differentiable and decreasing ($h'(X_0) < 0$). These assumptions imply, in particular, that more assimilation effort leads to higher costs, but a higher level of assimilation of the previous generation results in lower effort costs of assimilation for the current generation.

The purpose of this section is to understand the mechanism through which the assimilation effort will be transmitted between generations by taking X_0 as given.

In (2), α_i represents the weight placed by each individual i on her own income. This is the only ex ante heterogeneity that individuals have in this model. We assume that α_i differs among minority individuals and is drawn from a distribution over $[\underline{\alpha}, \overline{\alpha}]$; its cumulative distribution function (cdf) is given by $G(\alpha)$ and its density function by $g(\alpha)$. Thus, when an ethnic minority i decides

[8] This can be verified by showing that the cross effect of x_i and Y on the utility is negative, that is, $\frac{\partial^2 U}{\partial x_i \partial Y} < 0$.

[9] An alternative description of the group status could be modeled as $p(x_i)(Z - Y) + (1 - p(x_i))(Y - Z)$, where the first and second terms describe the relative group states that individual i receives from the viewpoint of the majority and the minority group, respectively. If x_i is large, her identity belongs more to the majority group, and hence, her utility decreases with Y. Since this model ignores the second term, only the effect of the group status through the first term prevails.

the level of her assimilation effort, she will primarily trade off a higher income against high effort costs. The perceived distance is the lowest at full or no assimilation and the highest at intermediate assimilation. The group status depends not only on the individual's choice but also on the other minority individuals' choice through average income Y. Finally, individual i's choice will also be affected by her α_i, that is, the weight she puts on her income in her utility function. Clearly, ethnic minority individuals with low (high) α_i will be less (more) likely to assimilate.

3.1.5 Analysis

Each minority individual maximizes $U_i := U(\alpha_i, x_i)$ with respect to x_i, taking the average income of minority individuals, Y, as given. Since X_0 is exogenous, here, we will not characterize the different possible equilibria. This will be done in Section 3.2 where we fully develop the dynamic model with X_0 endogenous. Here, we would like to investigate the different properties of our model when X_0 and Y are taken as given. We start with the following proposition.

Proposition 1

- (i) *A higher preference for own income α_i results in a higher degree of assimilation, x_i, a higher average income, Y, but a lower degree of assimilation for the other individuals x_j ($j \neq i$), that is, $\partial x_i/\partial \alpha_i > 0$, $\partial Y/\partial \alpha_i > 0$, $\partial x_j/\alpha_i < 0$ for all $i \neq j$.*

- (ii) *A higher average assimilation effort of the previous generation, X_0, results in a higher current average income, Y, and a higher assimilation effort for some individuals, that is, $\partial Y/\partial X_0 > 0$ and $\partial x_i/\partial X_0 > 0$ for some i.*

- (iii) *The effects of the weight of perceived distance, ρ, on individual i's assimilation effort, x_i, and the average income, Y, are ambiguous. However, a higher weight on the group status, σ, or a higher average income of majority individuals, Z, results in a higher average income, Y, and a higher assimilation effort, x_i, that is, $\partial Y/\partial \sigma \geq 0$, $\partial x_i/\partial \sigma \geq 0$, $\partial Y/\partial Z > 0$, and $\partial x_i/\partial Z > 0$ for all i.*

First (i), an increase in α_i leads to a higher incentive to assimilate for individual i and results in a higher average income Y. However, an increase in Y raises the social status of the minority group, which lowers the incentive to assimilate. The former effect dominates the latter for individual i, who thus assimilates more. However, other individuals assimilate less if their income increases. This implies that an ethnic minority individual i has a *higher* incentive to assimilate when other individuals belonging to the same ethnic group have low αs than when other individuals have high αs. It also implies that an individual with a higher α_i tends to assimilate more.

Second (ii), a higher average assimilation effort of the previous generation, X_0, decreases the effort costs to assimilate of the current generation. This yields a higher incentive to assimilate, which increases Y, the average income of the current generation. However, a higher Y implies a higher status of the minority group, which decreases their incentive to assimilate. Because of the heterogeneity in α_i, the former effect will dominate the latter for individuals with a large α_i whereas the opposite will hold true for individuals with a low α_i. Still, from the effect on Y, we know that a larger X_0 leads to a higher assimilation effort for at least some individuals. Thus, the impact of assimilation over generations depends on the effect of the group status over time. If the ethnic group has a high status, we will observe declines in assimilation over generations. If it is low, assimilation will increase over generations, which yields the possibility of multiple steady state equilibria. We will explore this issue when we extend the static model into a dynamic one (Section 3.2).

Finally (iii), the importance of the perceived distance, captured by ρ, has an ambiguous effect on assimilation because the perceived distance can be very low both when an minority individual totally rejects the majority's norm ($x_i = 0$) and when she totally assimilates to the majority's norm ($x_i = \overline{x}$). In contrast, a higher weight on the group status, σ, induces minority individuals to assimilate more, which leads to a higher average income for the ethnic minorities, Y. Finally, a higher average income of majority individuals, Z, increases the incentive to assimilate through a higher status of the majority group.

Because the effect through the group status is not taken into account when deciding their assimilation effort, we should expect that this effort's choice is inefficient from a welfare viewpoint.[10]

Let us determine the *optimal* assimilation effort and examine whether it differs from the *equilibrium* one. We consider the following standard Benthamite social welfare function:

$$W = \sum_{j=1}^{m} U(\alpha_j, x_j). \tag{3}$$

We can examine the efficiency of the equilibrium assimilation effort by checking the effects of an increase in x_i from equilibrium's value on social welfare (3). Such effects can be derived by evaluating the derivative of (3) with respect to x_i at the equilibrium level. If it is positive, the equilibrium assimilation effort is inefficiently low while, if it is negative, it is inefficiently high. This exercise is equivalent to comparing the equilibrium assimilation effort to the optimum that maximizes the social welfare, W.

Proposition 2 *The assimilation effort in any interior equilibrium is inefficiently high.*

[10] Note that this inefficiency does not arise from the assumption that each individual takes Y as given. Even if she internalizes a change of Y from her effort choice x_i, she does not internalize the effect of a change in x_i on the assimilation efforts of the other individuals through a change in Y. For the sake of simplicity, we assume that she ignores a change in Y when deciding on x_i and thus takes Y as given.

When a minority individual decides on her assimilation effort, she takes the average income, Y, as given. Hence, she ignores the effects of the change in Y caused by a change in x_i on her utility function. Moreover, she does not take into account the effect of such a change on the other individuals' utility. Indeed, an increase in x_i increases Y, which yields lower utility (i.e., lower gains from having empathy for the majority's relative status). Thus, the effects on her own and other individuals' utility work as negative externalities and make the assimilation effort inefficiently high.

3.2 Dynamic analysis

Let us now examine the dynamic process of our model, in which X_0 becomes endogenous.

3.2.1 Timing

Time is discrete, and each generation is economically active only during one period. At the beginning of each period t, m of new minority individuals enter the economy. Upon entering, each individual i decides on her assimilation effort, x_{it}, given the average assimilation effort of the previous generation, $X_{t-1} = (\sum_{j=1}^{m} x_{jt-1})/m$. The subscript t represents the period of entry of a generation. Then, the production takes place and the income, perceived distance, and group status of all ethnic monorities are realized; this determines the utility (2) of each individual i. At the end of the period, they exit the economy.

For tractability, we assume that the output per capita is given by $y(x_{it}) = y + \theta x_{it}$, the "empathy" function by $p(x_{it}) = x_{it}/\delta$, the assimilation cost function by $c(x_{it}) = \gamma x_{it}^2/2$, and the previous generation assimilation function by $h(X_{t-1}) = 1/(1 + X_{t-1})$, where y, θ, δ, and γ are all positive constants. Observe that y represents the income level with no assimilation effort, θ the marginal returns of the assimilation effort, and δ, the upper bound of the assimilation effort (i.e., $\overline{x} = \delta$), implying that $x_{it} \in [0, \delta]$, and γ is the cost parameter.

3.2.2 Symmetric steady-state equilibria

We first focus on the symmetric case where α_i is the same for all minority individuals, that is, $\alpha_i = \alpha$, $\forall i$. In this dynamic framework, a steady state equilibrium is given by a tuple (x_i^*, X^*, Y^*). The following proposition characterizes all possible stable steady-state equilibria.

Proposition 3 *Suppose that $\alpha_i = \alpha$ for all i and that (B.12) and (B.13) hold.*[11] *Then,*

[11] See the proof of Proposition 3 for a detailed analysis of the dynamics and what these conditions stand for. Condition (B.12) guarantees that there is a unique solution for the optimization problem of agent i (that is, the second-order condition is satisfied) while (B.13) ensures that X_t is always strictly positive. Indeed, condition (B.13) prevents an uninteresting equilibrium when minority individuals, on average, have no incentive to assimilate. To avoid such an extreme situation, we assume (B.13) throughout the rest of the paper.

- *when* $2\rho \leq \sigma\theta$*, there exists a unique, stable steady-state equilibrium;*
- *when* $2\rho > \sigma\theta$*, there exist multiple stable steady-state equilibria if there exist some* $\omega \in (0, \delta)$ *and* $\Theta \equiv \alpha\theta - \rho + \sigma(Z - y)/\delta$ *satisfying*

$$\left(1 + \omega + \omega^2/4\right)(2\rho - \sigma\theta)/\delta \geq \gamma$$

and

$$\gamma - (1 + 1/\delta)(2\rho - \sigma\theta) \leq \Theta \leq \gamma - (\omega/\delta + 1/\delta)(2\rho - \sigma\theta).$$

Indeed, the perceived distance, $p(x_i)(\overline{x} - x_i)$, is zero at $x_i = 0$ (no effort), since a minority individual has no empathy $p(0) = 0$ and is also zero at $x_i = \overline{x}$ (full effort), since there is no difference in effort with that of the majority. This implies that the perceived distance first increases then decreases as the minority assimilation effort level increases, which results in the convexity of the incentive to make an effort. Thus, since ρ and σ represent the utility weight on the perceived distance and group status, respectively, and θ the marginal returns of the assimilation effort, when ρ is sufficiently small (i.e., $2\rho \leq \sigma\theta$), the perceived distance effect's is less important compared to the role of the group status and cost, so that there is a unique, stable steady-state equilibrium. This is, in particular, due to the fact that the previous generation's assimilation effort is positively related to the current generation's assimilation effort through the cost reducing effect. When this condition does not hold (i.e. $2\rho > \sigma\theta$), the perceived distance's effect on incentive to make an effort prevails, yielding the possibility of multiple equilibria.

Adding to the convexity of the minority effort, the existence of multiple equilibria requires that Θ takes an intermediate value. It is the marginal return of the assimilation effort at $x_i = 0$; it is increasing in α (the preference for assimilation) and in $Z - y$ (the income gap between groups). When Θ lies in a particular interval, both the low-effort and the full-effort equilibrium coexist. The first inequality of Proposition 3 ensures the existence of the interval, and the second inequality requires Θ to lie in this interval. However, when it is very large (or very small, respectively), the minority individuals have a high (or low, respectively) incentive to assimilate. Hence, only the full-effort (or low-effort, respectively) equilibrium prevails.

Figure 1 depicts all the possible stable steady-state equilibria, where the solid curve represents the law of motion of X_t given by (B.15), while the dotted curve represents the possible stable steady-state equilibria. Figure 1(a) displays the case when $2\rho \leq \sigma\theta$ (i.e., the law of motion of X_t (B.15) is concave or linear), which consists of two sub-cases. In the left panel, there is a unique *interior* steady-state equilibrium whereas, in the right panel, there is a unique *corner* steady-state equilibrium, in which all minorities totally assimilate, that is, $x_i = \bar{x} = \delta$, for all i.

Figure 1(b) displays the case when $2\rho > \sigma\theta$[12], which consists of three sub-cases. In the upper-left panel, there is a unique interior equilibrium. In the upper-right panel, there is a unique corner equilibrium in which all minorities totally assimilate. Finally, in lower panel, multiple equilibria exist. Hence, as stated in Proposition 3, in order for multiple equilibria to emerge, the inequality that $2\rho > \sigma\theta$ must hold true.

Proposition 4 *Suppose that there is an interior, stable steady-state equilibrium X^*. Then, an increase in $Z - y$ or α, or a decrease in γ, increases X^*. The effects of an increase in θ, ρ, σ, or δ on X^* are ambiguous.*

As the income inequality $Z - y$ between a majority individual and a minority individual with no assimilation effort increases, the gains from having a higher empathy to majority increase due to changes in the group status. A larger α implies that a minority individual puts more weight on income, inducing a larger assimilation effort. A smaller γ makes assimilation less costly, encouraging a minority individual to make more effort. The effects of other parameters are ambiguous because of the convexity of the perceived distance's effect on the effort incentive.

Consider the following welfare function: $TW = \sum_{t=0}^{\infty} \beta^t W_t$, where $0 < \beta < 1$ is the discount factor. We have the following result:

Proposition 5 *The assimilation effort in the interior equilibrium can be inefficiently low or high.*

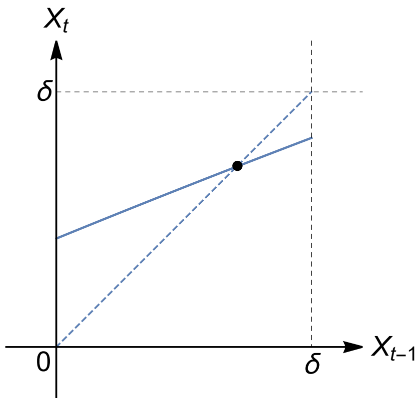
Indeed, in terms of assimilation, each individual effort exerts *negative externalities* on the other individuals of the *same* generation, since more assimilation decreases the gains from having empathy (or conformity) for the majority's relative status, but *positive externalities* to the *next* generation, since more assimilation effort decreases the assimilation cost of the next generation. Therefore, the effect of assimilation is negative in the *short run* while positive in the *long run*. This is why the net effect at the steady-state equilibrium is ambiguous and depends on the value of β (the discount rate).

3.2.3 Asymmetric steady-state equilibria

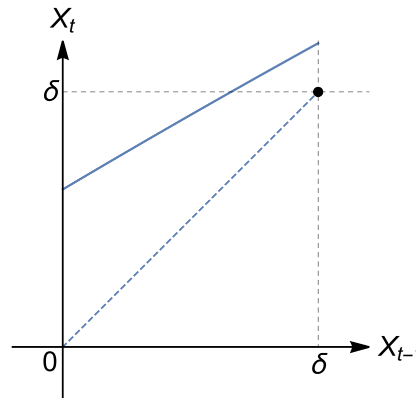
Now, consider the case when agents are ex ante heterogenous, that is, individuals have different α_i. When α_i differs among individuals, we may have corner solutions, that is, $x_i = 0$ or $x_i = \delta$, for some individuals as well as interior solutions for others, depending on the value of their αs. For a large value of X_{t-1}, the first-order condition (B.14) can make x_i larger than δ for individuals with high α_i, which results in a corner solution $x_i = \delta$. For a small value of X_{t-1}, the first-order

[12] When $2\rho > \sigma\theta$, we have $d^2X_t/dX_{t-1}^2 > 0$, which means that the law of motion of X_t (B.15) is convex.

Figure 1: Steady-state equilibria when all agents have the same α

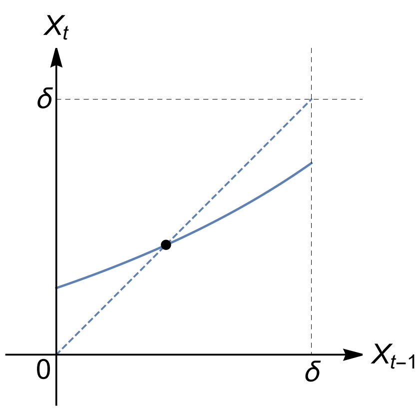


Unique interior equilibrium

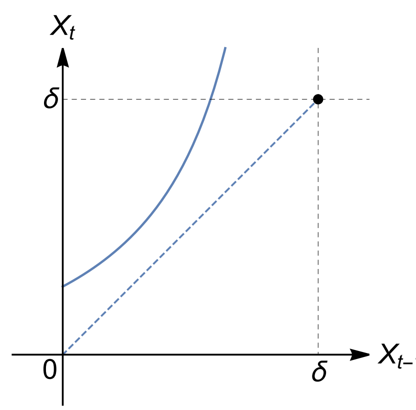


Unique corner equilibrium

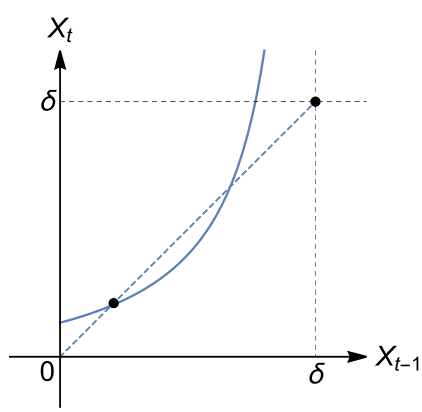
(a) The case of $2\rho \leq \sigma\theta$



Unique interior equilibrium



Unique corner equilibrium



Multiple equilibria

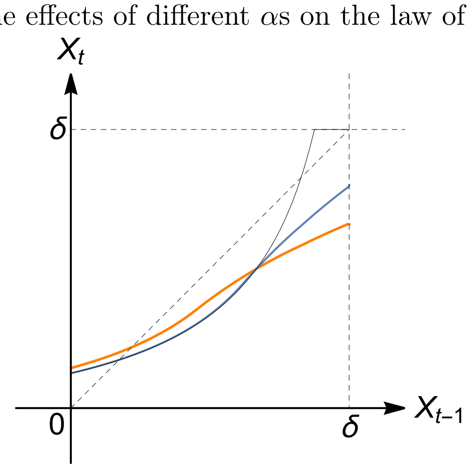
(b) The case of $2\rho \geq \sigma\theta$

condition (B.14) can make x_i smaller than 0 for individuals with low α_i, which results in a corner solution, $x_i = 0$. Hence, heterogeneity in α_i makes the curve of (B.15) in the $X_{t-1} - X_t$ plan flatter, which decreases the possibility of multiple equilibria. This is confirmed for the numerical example displayed in Figure 2.[13]

Indeed, Figure 2 displays three curves where each plots (B.15) for different values of α_i: a thin black curve, which represents the symmetric case where $\alpha_i = 2$ for all minority individuals; a blue curve, which represents the case where the variance of α_i is 0.34 while the mean is fixed; and an orange curve, which represents the case where the variance of α_i is 1.37 while the mean is fixed. All the other parameters are the same for the three curves.[14]

Starting from the symmetric case (i.e., the thin black curve), we consider a mean preserving spread on the distribution of α_i. First, we consider the case when $\alpha_i = 2 - 0.02i$, for $i = 1, .., 50$ and $\alpha_i = 2 + 0.02(i - 50)$ for $i = 51, .., 100$, resulting in a mean of $\bar{\alpha} = 2$ and a variance of $var(\alpha_i) = 0.34$. This case is displayed by the blue curve. We then consider the case when $\alpha_i = 2 - 0.04i$, for $i = 1, .., 50$ and $\alpha_i = 2 + 0.04(i - 50)$ for $i = 51, .., 100$, resulting in a mean of $\bar{\alpha} = 2$ and a variance of $var(\alpha_i) = 1.37$. This case is displayed by the orange curve. In Figure 2, we can see that a larger variance in α_i makes the curve of (B.15) flatter. This implies that, when there are multiple equilibria in the symmetric case (same α for all individuals), there is a unique interior equilibrium in the asymmetric case.

Figure 2: The effects of different αs on the law of motion of X



[13]Since the dynamic asymmetric model becomes very complicated, we do not have a total characterization of all stable steady-state equilibria as in the symmetric case. However, we can understand how different αs affect the equilibria.

[14]In Figure 2, we set $m = 100$, $Z = 3$, $y = 1$, $\rho = 4/3$, $\sigma = 1/2$, $\theta = 1/3$, $\delta = 1$, and $\gamma = 5$.

4 Majority's decision

Let us now introduce the majority's decision to the model and their decision whether to accept the minority's assimilation.

4.1 Static analysis

Consider first the static framework. Suppose there are n majority individuals, with $n > m$. Majority individuals do not need to assimilate, and hence, we assume their income is fixed at Z despite heterogeneity in the preference for material returns, η_k ($k = 1, 2, .., n$). This also implies that their social status is fixed, and we denote the returns from it with B. However, majority individuals incur a disutility from the perceived distance between their $\bar{x}$ and X, the average assimilation effort of the ethnic minorities. One can think, for example, of language and the cost of interaction with non-fluent individuals. Indeed, for a majority individual, the disutility from the perceived distance is given by $(\lambda - e_k)q(X)(\overline{x} - X)$, where λ represents full acceptance and $0 \leq e_k \leq \lambda$ the key decision variable, which is the majority's degree of accepting the minority individuals, while $q(X)$ describes the majority's possibility of communicating with minority individuals. We assume that $q'(X) > 0$, $q''(X) \leq 0$ and $q(0) = 0$, implying that the majority individuals have more opportunities to communicate with the minority individuals the more the latter assimilate. When minorities totally reject assimilation ($X = 0$), then there is no communication between the majority and minority individuals. Moreover, given the degree of acceptance, e_k, the cost of the perceived distance increases with $q(X)$, the possibility of communicating with ethnic minorities, and with $\overline{x} - X$, the perceived distance. These features are motivated by Colussi et al. (2021), who show that majority individuals' attitudes toward a minority are worsened by the salience of minority people and their cultural dissimilarities.

We assume that the majority individuals, who choose e_k, their *degree of tolerance* toward the minority group, bear some effort cost $d(e_k)$, with $d'(e_k) > 0$ and $d''(e_k) > 0$. Summarizing our discussion, we can write the utility of a majority individual k as follows:

$$V_k = \eta_k Z + B - (\lambda - e_k)q(X)(\overline{x} - X) - d(e_k). \tag{4}$$

For our model to be consistent, we need to introduce e_k into the utility function (2) of a minority individual. For that, we assume that the effort cost of attaining a particular assimilation effort x_i becomes lower, the higher the average value is of e_k. Indeed, when the majority individuals are more willing to accept minority individuals in their country, the cost of assimilation becomes lower. By letting E denote the majority individuals' average acceptance, that is, $E = \sum_{i=1}^{n} e_k/n$, we can write the utility (2) of a minority individual as follows:

$$U_i = \alpha_i y(x_i) - \rho p(x_i)(\overline{x} - x_i) + \sigma p(x_i)(Z - Y) - s(E)h(X_0)c(x_i), \tag{5}$$

where $s(E)$ is decreasing in E.

To summarize, the key aspect of our model is that the minority's average assimilation effort $X = (\sum_{j=1}^{m} x_j)/m$ affects e_k the majority's tolerance decision toward minorities, which, in turn, through E, the majority individuals' average acceptance, affects the individual decision to assimilate x_i, and, thus, X.

Each majority individual chooses e_k that maximizes (4) by taking X as given. We obtain:[15]

$$q(X)(\overline{x} - X) = d'(e_k^*). \tag{6}$$

This shows the tradeoff faced by majority individuals when choosing e_k. On the one hand, the higher e_k, the lower their perceived distance, which increases their marginal utility. On the other, higher e_k increases the costs of exerting effort e_k, which decreases their marginal utility.

It is straightforward to evaluate the effect of the average degree of minority assimilation, X, on the degree of majority acceptance, e_k^*.

Proposition 6 *When the average degree of minority assimilation, X, is small, an increase in X increases the degree of majority acceptance, e_k^*, as well as its average, E^*. On the contrary, when X is large, an increase in X decreases e_k^* and E^*.*

This proposition shows that there is an inverted $U-$shaped relationship between e_k, the degree of tolerance of a majority individual k, and X, the average level of assimilation in the society. Indeed, when X is close to zero, that is, no assimilation in the society, then increasing X leads to a higher e_k, since there is no possible communication between minority and majority individuals ($q(0) = 0$) and thus the majority is willing to increase e_k when X increases. Conversely, when X is close to its maximum value, $\bar{x}$, there is perfect communication between minority and majority individuals, since they have the same "assimilation" level, and the majority group is not ready to pay an extra cost to increase e_k. Thus, there is an intermediary value of X, denoted by $0 < \tilde{X} < \bar{x}$, such that for all $X \leq \tilde{X}$, increasing X increases e_k and for all $X > \tilde{X}$, increasing X decreases e_k.

This is what we empirically showed in Section 2. Using the 2019 American Community Survey, we observed that an increase in the minority's assimilation rate raised the majority's degree of acceptance, but this relationship was not linear (figures A1 and A2 in Appendix A). In particular, in these figures, we have an inverted $U-$shaped relationship between intermarriage rates (capturing

[15] The second-order condition is always satisfied.

the majority's tolerance toward the minority) and English proficiency or naturalization rate (both a measure of the minority's assimilation).

Obviously, X is endogenous and chosen by the minority individuals. The next section investigates these effects when both e_k and x_i are choice variables.

4.2 Dynamic analysis

Let us now consider a dynamic framework where X_0 is endogenous. For tractability, we assume $q(X_t) = \varepsilon X_t$, $d(e_{kt}) = \mu e_{kt}^2/2$, $y(x_{it}) = y + \theta x_{it}$, and $s(E_t) = \lambda - E_t$. Under these specifications, we can solve the minority's and majority's optimization and obtain

$$e_{kt}^* = E_t^* = \frac{\varepsilon X_t(\delta - X_t)}{\mu}, \quad (7)$$
$$x_{it}^* = \frac{\alpha_i\theta - \rho + \frac{\sigma}{\delta}(Z - Y_t)}{\frac{\gamma(\lambda - E_t)}{1+X_{t-1}} - \frac{2\rho}{\delta}},$$

which yields the following law of motion:

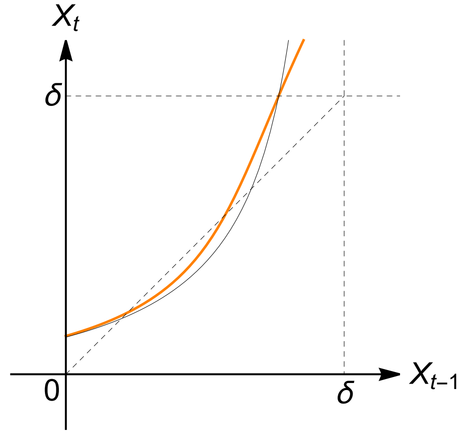
$$X_t = \frac{A\theta - \rho + \frac{\sigma}{\delta}(Z - y)}{\frac{\gamma(\lambda - E_t)}{1+X_{t-1}} + \frac{\sigma\theta - 2\rho}{\delta}}. \quad (8)$$

If we fix the majority's acceptance, E_t, the law of motion of X_t, (8) has exactly the same properties as the one shown in Proposition 4 (where the majority's choice was not modeled). However, as we can see from (7), the majority's decision adds another non-linearity, preventing us from analytically characterizing the different possible steady-state equilibria.

Thus, to analyze (8), we numerically plot in Figure 3 the law of motion of X (i.e., equation (8)), where the solid, thin black curve represents the symmetric case *without majority acceptance* while the orange curve represents the symmetric case *with majority acceptance*.[16] We can see from this figure that the curve becomes steeper for smaller values of X_{t-1} and flatter for large values of X_{t-1}. This implies that the introduction of a majority acceptance decision weakens the convexity of the minority assimilation, reducing the possibility of multiple equilibria.

[16] In Figure 3, we consider the symmetric case (i.e., $\alpha_i = \alpha$ for all i), and we set $Z = 3$, $y = 1$, $\rho = 4/3$, $\sigma = 1/2$, $\theta = 1/3$, $\delta = 1$, $\gamma = 5$, $\varepsilon = 1/2$, $\mu = 4$, and $\lambda = 1$.

Figure 3: The effects of majority acceptance on the law of motion of X



Let us now consider welfare issues. We have the following result:

Proposition 7 *Consider a welfare function* $TW_a = \sum_{t=0}^{\infty} \beta^t \left(\sum_{j=1}^{m} U_{jt} + \sum_{k=1}^{n} V_{kt} \right)$ *that incorporates both the minority's and majority's utility function.*

- *In any interior equilibrium, the majority's degree of acceptance is inefficiently low while the minority's assimilation effort can be inefficiently low or high.*

- *Suppose that* α *is the same for all minority individuals (symmetric case). Then, if* $X_t < \delta/2$ *(respectively,* $X_t \geq \delta/2$*), the assimilation effort is inefficiently high (respectively, low) for a sufficiently large number of majority individuals relative to minority individuals,* n/m*, and is inefficiently low (respectively, high) for a sufficiently small* n/m *and a sufficiently large (respectively, small)* β*.*

As we can see from the minority's utility function (5), a higher majority acceptance yields a lower effort cost for minority individuals. Hence, each majority individual acceptance generates *positive externalities* for ethnic minorities, making it inefficiently low. In contrast, each minority individual effort exerts *negative externalities* on the other minority individuals of the same generation but *positive externalities* on the next generation. Furthermore, it affects the majority individuals differently depending on the minority assimilation level, X_t. If it is sufficiently low and satisfies $X_t < \delta/2$, the minority assimilation increases the majority's disutility from the perceived distance, yielding negative externalities. If $X_t \geq \delta/2$, it decreases the majority's disutility from the perceived distance, yielding positive externalities.

The number of majority individuals relative to minority individuals, n/m, determines the strength of these externalities on majority individuals. If n/m is large, the externalities on majority individuals dominate the externalities within minority individuals. If n/m is small, then the overall externalities' sign depends on the externalities within minority individuals. In such a case, β, the discount rate, determines the relative strength of the positive externalities on the next generation to the negative externalities on the current minority individuals.

4.3 Policy implications

In this framework, a key policy is how to reach a "harmonious" society that integrates the majority and minority individuals. For that, there are two possible ways: either the planner enhances the minority's assimilation effort (for example, by providing incentives to learn the native language),[17] or it increases the majority's acceptance effort (for example, by geographically mixing majority and minority individuals).[18]

Let us assume that the planner controls either the minority's assimilation effort or the majority's acceptance, while the other group's behavior is left uncontroled and determined endogenously. Let us assume a symmetry in effort within each group (same α and same η). From equations (7) and (8), the long-run equilibrium conditions under this policy intervention are described as follows.

$$X^* = \frac{A\theta - \rho + \frac{\sigma}{\delta}(Z - y)}{\frac{\gamma(\lambda - \overline{E})}{1 + X^*} + \frac{\sigma\theta - 2\rho}{\delta}} \equiv X^*(X^*, \overline{E}), \tag{9}$$

$$E^* = \frac{\varepsilon \overline{X}(\delta - \overline{X})}{\mu} \equiv E^*(\overline{X}). \tag{10}$$

Equation (9) corresponds to the case where the government only controls the majority's acceptance effort (that is, it determines $\overline{E}$), while X^* is determined by the equilibrium condition. Equation (10) corresponds to the case where the government only controls the minority assimilation effort (that is, it determines $\overline{X}$), while E^* is determined by the equilibrium condition. We have the following result:

Proposition 8 *Assume that the planner decides upon either the majority's acceptance effort or the minority's assimilation effort while the behavior of the other side is determined by the market.*

- *An increase in the majority's degree of acceptance, $\overline{E}$, raises the minority's assimilation effort, X^*.*

[17] For an overview of language policies and programs in several countries in Europe, Australia, and North America, see Li and Sah (2019).

[18] In the United States, the aim of the Moving to Opportunity Programs (MTOs) is precisely to integrate geographically different segments of the population, often minority and majority families. In the MTOs, families in areas with high unemployment and poverty are given the possibility of moving to areas with a higher level of gainful employment and education as well as better schools and education. See, e.g., Chetty et al. (2016).

- *An increase in the minority's assimilation effort, $\overline{X}$, increases the majority's degree of acceptance, E^*, if and only if $\overline{X}$ is smaller than $\delta/2$.*

Proposition 8 shows that an increase in $\overline{E}$ always increases X^*. This means that the more the majority individuals are tolerant toward the minority group, the more the minority individuals will assimilate to the majority group. However, an increase in $\overline{X}$ does not always increase E^*, especially if assimilation is already very high in the society. Therefore, from the viewpoint of integration, if the government can control only one side, it should focus on the majority group.

Proposition 9 *Starting from an interior equilibrium, the policy that either increases the minority's assimilation effort X^* or the majority's degree of acceptance E^* has the following effects:*

- *Both policies increase the welfare of the majority group for any generation if and only if X^* is larger than $\delta/2$ and decrease it otherwise.*

- *The policy that increases X^* decreases the welfare of the minority group of the current generation if X^* is larger than $\delta/2$, while it has an ambiguous effect for the minority of any other generation.*

Indeed, when X^* is larger than $\delta/2$, the minority assimilation decision generates positive externalities on the majority individuals (since it decreases the perceived distance). Hence, the government's optimal intervention is to raise X^* to improve the majority's welfare. The policy that controls E^* has similar effects, since it also increases X^*.

In contrast, the effects of the policies on the minority's welfare are almost always ambiguous because of various effects. Indeed, the minority assimilation effort creates negative externalities on the other minority individuals of the same generation (since it reduces the status difference $Z - Y$) but positive externalities on the next generation (since it reduces the assimilation costs). Furthermore, both policies may increase or decrease the majority's acceptance rate, which brings positive or negative externalities to the minority group of the other generations by affecting their assimilation cost.

Note that the policy effects on the minority's welfare differ across generations, even though the effect on the majority's welfare is common to all generations. This is because, when the policy starts, the first generation cannot benefit from the positive inter-generational externalities. Thus, the policy that controls the minority's assimilation hurts the welfare of the current minority group by decreasing the majority's acceptance when the degree of assimilation is already high.

To summarize, both policies have a similar positive effect on the majority's welfare when X^* is large enough. However, the impact on the minority's welfare is ambiguous and differs across generations and types of policies. In particular, when X^* is large, although it may benefit future generations, the policy intervention on the minority's assimilation decreases the majority's acceptance and negatively affects the welfare of the minority of the current generation.

5 Conclusion

In this paper, we advance the role of the majority group in the assimilation of ethnic minorities and how they interact with each other. To the best of our knowledge, at least from a theoretical perspective, research has neglected to examine how the majority group members' orientations affect minorities' assimilation strategies, that is, how the perceptions of the majority group members affect how ethnic minorities should adapt to the dominant culture within the host country. This line of research is important because the dominant group of society plays a powerful role in shaping minorities' assimilation/acculturation strategies (Berry, 2003). For example, the majority group members might prevent ethnic groups from fully participating in society if they possess negative stereotypes about these groups or if they consider these groups' economic and social status within the host country to be adverse to the dominant group members (Sayegh and Lasry, 1993). Thus, the potential obstacles to the social interactions of ethnic minorities with host majority group members need to be investigated in order to properly assess and examine the assimilation of the minority group.

This is what we do in the current study, by showing that the assimilation decisions of the minority group may not only affect the well-being of ethnic minority individuals but also that of the majority individuals. The model also focuses on investigating the assimilation strategies adopted by minorities in the host community and the relational outcomes that are the product of both minority and majority groups' attitudes regarding each other. Therefore, the model suggests that the combinations of majority group members' tolerance toward minorities and minorities' assimilation strategies could produce consensual outcomes between these two groups.[19]

In the policy implications of the model, we show that a well-assimilated minority group generates positive externalities for the majority group, which increases the majority's welfare. Thus, the planner wants to increase assimilation in the country. However, this does not necessary benefit the current generation's minority group because it reduces their relative status in the society but increases the welfare of the next generation by reducing their assimilation costs.

With the recent refugee crisis in Africa and the Middle East, we believe that the two-way interaction between the minority's assimilation decision and the majority's tolerance of the minority group highlighted in this paper needs to be incorporated by the European and American governments to design a successful policy for the assimilation and integration of ethnic minorities and immigrants

[19] For example, Bourhis et al. (1997) have argued that the larger the differences between the attitudes of majority group members and those of minorities, the more conflicting their relationship will be.

in their countries.

References

Akerlof, G. (1997). Social distance and social decisions. *Econometrica 65*, 1005–1027.

Akerlof, G. and R. Kranton (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton: Princeton University Press.

Algan, Y., C. Dustmann, A. Glitz, and A. Manning (2010). The economic situation of first- and second-generation immigrants in France, Germany, and the UK. *Economic Journal 120*(542), F4–F30.

Arai, M. and P. Skogman Thoursie (2009). Renouncing personal names: An empirical examination of surname change and earnings. *Journal of Labor Economics 27*, 127–147.

Arendt, J., I. Bolvig, M. Foged, L. Hasager, and G. Peri (2020). Language training and refugees' integration. NBER Working Paper No. 26834.

Benabou, R. and J. Tirole (2011). Identity, morals, and taboos: Beliefs as assets. *Quarterly Journal of Economics 126*, 805–855.

Berry, J. (2003). Conceptual approaches to acculturation. In: K.M. Chun, P.B. Organista and G. Marin (Eds.)*, Acculturation: Advances in Theory, Measurement, and Applied Research,* Washington, DC: American Psychological Association, 17–37.

Bezin, E. and F. Moizeau (2017). Cultural dynamics, social mobility and urban segregation. *Journal of Urban Economics 99*, 173–187.

Biavaschi, C., C. Giulietti, and Z. Siddique (2017). The economic payoff of name Americanization. *Journal of Labor Economics 35*, 1089–1116.

Biavaschi, C., C. Giulietti, and Y. Zenou (2021). Social networks and political assimilation in the Age of Mass Migration. CEPR Discussion Paper No. 16182.

Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou (2011). Formation and persistence of oppositional identities. *European Economic Review 55*, 1046–1071.

Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou (2016). Bend it like Beckham. Ethnic identity and integration. *European Economic Review 90*, 146–164.

Borjas, G. (1985). Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics 3*(4), 463–489.

Borjas, G. (1987). Self-selection and the earnings of immigrants. *American Economic Review 77*, 531–553.

Borjas, G. (1992). Ethnic capital and intergenerational mobility. *Quarterly Journal of Economics 107*, 123–150.

Boucher, V. (2016). Conformism and self-selection in social networks. *Journal of Public Economics 136*, 30–44.

Boucher, V., T. S., M. Vlassopoulos, J. Wahba, and Y. Zenou (2021). Ethnic mixing in early childhood: Evidence from a randomized field experiment and a structural model. IZA Discussion Paper No. 14260.

Bourhis, R., L. Moise, S. Perreault, and S. Senecal (1997). Towards an interactive acculturation model: A social psychological approach. *International Journal of Psychology 32*, 369–386.

Carvalho, J.-P. (2016). Identity-based organizations. *American Economic Review, Papers and Proceedings 106*(5), 410–414.

Chetty, R., N. Hendren, and L. F. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity Experiment. *American Economic Review 106*(4), 855–902.

Chiswick, B. and P. Miller (1995). The endogeneity between language and earnings: International analyses. *Journal of Labor Economics 13*, 246–288.

Chiswick, B. and P. Miller (1996). Ethnic networks and language proficiency among immigrants. *Journal of Population Economics 9*, 19–35.

Chiswick, B. and P. Miller (2005). Linguistic distance: A quantitative measure of the distance between english and other languages. *Journal of Multilingual and Multicultural Development 26*, 1–11.

Chiswick, B. and P. Miller (2011). The 'negative' assimilation of immigrants: A special case. *Industrial and Labor Relations Review 64*, 502–525.

Colussi, T., I. Isphording, and N. Pestel (2021). Minority salience and political extremism. *American Economic Journal: Applied Economics 13*, 237–271.

Damm, A. (2009). Ethnic enclaves and immigrant labor market outcomes: Quasiexperimental evidence. *Journal of Labor Economics 27*(2), 281–314.

de la Rica, S., A. Glitz, and F. Ortega (2015). Immigration in Europe: Trends, policies, and empirical evidence. In: Barry R. Chiswick and Paul W. Miller (Eds.)*, Handbook of the Economics of International Migration,* Amsterdam: Elsevier Science, 1303–1362.

Drouhot, L. G. and V. Nee (2019). Assimilation and the second generation in Europe and America: Blending and segregating social dynamics between immigrants and natives. *Annual Review of Sociology 45*, 177–199.

Dustmann, C. and F. Fabbri (2003). Language proficiency and labour market performance of immigrants in the UK. *Economic Journal 113*, 695–717.

Dustmann, C., K. Vasiljeva, and A. Damm (2019). Refugee migration and electoral outcomes. *Review of Economic Studies 86*, 2035–2091.

Edin, P., P. Fredriksson, and O. Åslund (2003). Ethnic enclaves and the economic success of immigrants. Evidence from a natural experiment. *Quarterly Journal of Economics 118*, 329–357.

Feenstra, R., R. Inklaar, and P. Timmer (2015). The next generation of the Penn World Table. *American Economic Review 105*(10), 3150–3182.

Hammarstedt, M. (2009). Intergenerational mobility and the earnings position of first-, second-, and third-generation immigrants. *Kyklos 62*(2), 275–292.

Hangartner, D., E. Dinas, M. Marbach, K. Matakos, and D. Xefteris (2019). Does exposure to the refugee crisis make natives more hostile? *American Political Science Review 113*(2), 442–455.

Hauk, E. and H. Mueller (2015). Cultural leaders and the clash of civilizations. *Journal of Conflict Resolution 59*, 367–400.

Kim, Y.-C. and G. Loury (2019). To be, or not to be: Stereotypes, identity choice and group inequality. *Journal of Public Economics 174*, 36–52.

Kónya, I. (2007). Optimal immigration and cultural assimilation. *Journal of Labor Economics 25*, 367–391.

Lazear, E. (1999). Culture and language. *Journal of Political Economy 107*, S95–S126.

Li, G. and P. K. Sah (2019). Immigrant and refugee language policies, programs, and practices in an era of change. Promises, contradictions, and possibilities. In: Steven J. Gold and Stephanie J.

Nawyn (Eds.), *Routledge International Handbook of Migration Studies, 2nd Edition,* New York: Routledge, 325–338.

Li, Q. (2013). Language and urban labor market segmentation: Theory and evidence. *Journal of Urban Economics 74*, 27–46.

Liu, X., E. Patacchini, and Y. Zenou (2014). Endogenous peer effects: Local aggregate or local average? *Journal of Economic Behavior and Organization 103*, 39–59.

Maskileyson, D., M. Semyonov, and E. Davidov (2021). Intergenerational mobility and the earnings position of first-, second-, and third-generation immigrants. *Population, Space and Place 27*(6), e2426.

Meng, G. and R. Gregory (2005). Intermarriage and the economic assimilation of immigrants. *Journal of Labor Economics 23*, 135–175.

Moriconi, S., G. Peri, and R. Turati (2019). Immigration and voting for redistribution: Evidence from European elections. *Labour Economics 61*, 101765.

Munoz-Herrera, M. (2021). Identity change and the cost of social division: Theory and experimental evidence. Unpublished manuscript, New York University Abu Dhabi.

Nosofsky, R. (1986). Attention, similarity and the identification-categorization relationship. *Journal of Experimental Psychology: General 115*, 39–57.

Panebianco, F. (2014). Socialization networks and the transmission of interethnic attitudes. *Journal of Economic Theory 150*, 583–610.

Patacchini, E. and Y. Zenou (2012). Juvenile delinquency and conformism. *Journal of Law, Economic, and Organization 28*, 1–31.

Prummer, A. and J.-P. Siedlarek (2017). Community leaders and the preservation of cultural traits. *Journal of Economic Theory 168*, 143–176.

Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2020). IPUMS USA: Version 10.0. Minneapolis, MN: IPUMS, https://www.ipums.org/.

Sambanis, N. and M. Shayo (2013). Social identification and ethnic conflict. *American Political Science Review 107*, 294–325.

Sato, Y. and Y. Zenou (2020). Assimilation patterns in cities. *European Economic Review 129*, 103563.

Sayegh, L. and J. Lasry (1993). Immigrants' adaptation in Canada: Assimilation acculturation, and orthogonal cultural identification. *Canadian Psychology 34*, 98–109.

Shayo, M. (2009). A model of social identity with an application to political economy: Nation, class, and redistribution. *American Political Science Review 103*, 147–174.

Shayo, M. (2020). Social identity and economic policy. *Annual Review of Economics 12*, 355–389.

Stark, O., M. Jakubek, and K. Szczygielski (2018). Community cohesion and assimilation equilibria. *Journal of Urban Economics 107*, 79–88.

Steinmayr, A. (2021). Contact versus exposure: Refugee presence and voting for the far-right. *Review of Economics and Statistics,* forthcoming.

Tabellini, M. (2020). Gifts of the immigrants, woes of the natives: Lessons from the Age of Mass Migration. *Review of Economic Studies 87*, 454–486.

Turner, R., M. Hogg, P. Oakes, S. Reicher, and M. Wetherell (2010). *Rediscovering the Social Group: A Self-Categorization Theory.* Oxford: Blackwell Publishing.

Ushchev, P. and Y. Zenou (2020). Social norms in networks. *Journal of Economic Theory 185*, 104969.

Verdier, T. and Y. Zenou (2017). The role of social networks in cultural assimilation. *Journal of Urban Economics 97*, 15–39.

Verdier, T. and Y. Zenou (2018). Cultural leaders and the dynamics of assimilation. *Journal of Economic Theory 175*, 374–414.

Xu, D. (2017). Acculturational homophily. *Economics of Education Review 59*, 29–42.

Appendix

A Stylized facts

In this appendix, we provide some descriptive evidence (no causality is claimed) of the predictions on which we focus in the theoretical model.

A.1 What matters for the assimilation of immigrants?

As is usually the case, assimilation is measured by the English proficiency of immigrants. Here, we estimate a probit model by maximum likelihood estimation in order to examine whether we may observe predictions regarding the assimilation of immigrants.

A.1.1 Data

The main data source is the 2019 American Community Survey (ACS) supplied by the Integrated Public Use Microdata Series (IPUMS) (Ruggles et al., 2020). The estimation is performed for ethnic minority people (including the immigrants) who are naturalized citizens, those who are not citizens, and those who are not citizens but have received their first papers. Moreover, we focus on the adult working population, that is, individuals who are at least 18 years old and at most 64 years old. Among the individuals whose English proficiency information is available, we identified those who *speak English fluently* and construct a dummy for it (*EnPro*), which is the *dependent variable* of our probit estimation. Because of the availability of language distance data, the sample size for the baseline estimation is 172,303. Moreover, because of the availability of souce country variables, the sample size for the case with all control variables is 163,558.

We focus on the following two *independent variables*. First, we focus on *PastAvgEnPro*, which is the past share of immigrants who speak English fluently. This variable is available by county and by language spoken at home calculated from the 2010 American Community Survey. In our model, this variable is captured by the past average assimilation degree (X_{t-1}), and its estimated coefficient is expected to have a positive sign.

Second, we focus on *LangDist*, which is the linguistic distance from English. Chiswick and Miller (2005) developed an index of linguistic similarity, where the reference language is English. The index takes values between 1 and 3, with a larger value representing a greater linguistic similarity with English. We take its inverse, which is interpreted as the *linguistic distance* from English of each language spoken at home. We assign a 0 to those who speak English at home. This variable corresponds to the cost of assimilation effort in our model. Its estimated coefficient is expected to have a negative sign.

We add different controls; their description can be found in Table A1. Observe that we include control variables of the *source countries* determined by birthplace information. These source country variables come from the Penn World Table 10 (Feenstra et al., 2015). As shown by Kónya (2007), the relationship between the host and source countries can significantly affect the assimilation behavior of ethnic minorities. Summary statistics of all variables are given in Table A1.

Table A1: Descriptive Summary Statistics of the American Community Survey dataset

Variables	Description	Mean	S.D.
EnPro	English proficiency dummy (=1 if an immigrant speaks English fluently)	0.792	-
InterMarriage	Interracial marriage dummy (=1 if a person is interracially married)	0.098	-
Naturalization	Naturalization dummy (=1 if an immigrant is naturalized)	0.529	-
PastAvgEnPro	Past share of immigrants who speak English fluently by county and by language spoken at home for the year 2010	0.718	0.223
LangDist	Linguistic distance from English	0.407	0.236
NativeIncome	County-median income of those who speak only English (US dollars, in natural logarithm)	11.4	8.8
FemaleDummy	Gender dummy (=1 if a person is female)	0.521	-
Age	Age	46.8	11.0
SchoolYear	Schooling years	13.3	2.90
AgeImmig	Age of immigration	22.5	11.9
YearImmig	Years after immigration	20.9	13.1
MADummy	Metropolitan Area (MA) dummy (=1 if a person locates in a MA)	0.997	-
RelativeGDP	Real GDP per capita of the source countty relative to that of the US	0.219	0.003
RelativePop	Population of the source countty relative to that of the US	0.801	1.41
LangShare	Share of the language spoken at home among immigrants	0.174	0.266
AvgEnPro	Share of immigrants who speak fluently English by county	0.818	0.084
AvgNatu	Share of naturalized immigrants by county	0.525	0.093

Notes: Variables other than PastAvgEnPro are for the year 2019. PastAvgEnPro is for the year 2010.

A.1.2 Results

We estimate the following model:

$$EnPro_{i,c} = \beta_0 + \beta_1 PastAvgEnPro_c + \beta_2 LangDist_{i,c} + Controls + \epsilon_{i,c},$$

where the subscript i, c denotes an individual i in county c. In addition, $\epsilon_{i,c}$ is the error term and $Enpro_{i,c}$ is a dummy variable:

$$EnPro_{i,c} = \begin{cases} 1 & \text{if an immigrant speaks English fluently} \\ 0 & \text{otherwise} \end{cases}$$

Table A2 displays the probit estimation results, while Table A3 shows the marginal effects on the probability that $Enpro_{i,c} = 1$. We see that the past average assimilation (*PastAvgEnPro*), that is, the past share of immigrants who speak English fluently, and the linguistic distance (*LangDist*), are both significant and have a positive and a negative sign, respectively. Indeed, the past share of immigrants who speak English fluently is positively correlated with the probability of speaking fluently English today. The lower the linguistic distance from English, the higher the chance of being assimilated (i.e., speaking English fluently).

A.2 Does the degree of native individuals' acceptance matter for the assimilation of immigrants?

We would like to examine the relationship between the assimilation of immigrants and the majority's acceptance of the minority culture. To measure assimilation, we use the share of immigrants who speak *English fluently* by county or the share of *naturalized immigrants* by county. To measure the degree of native acceptance, we use the *interracial marriage of non-Hispanic whites.* In order to capture the possibly non-monotonic relationship predicted by our model, we include a quadratic term.

A.2.1 Data

The main data source is again the 2019 American Community Survey supplied by IPUMS. For a native population, we focus on the non-Hispanic white adult working population, that is, those who are non-Hispanic white, at least 18 years old, and at most 64 years old. The American Community Survey provides each person's detailed race information. We construct an interracial marriage dummy, which takes 1 if a white person marries someone from another race. This is the proxy for our model's acceptance of majority individuals, e_t, which is the dependent variable in our estimations and is denoted by $InterMarriage_{i,c}$. For this analysis, the estimation sample size is 368,881.

The main independent variable $EnPro_c$ is the fraction of immigrants who speak English fluently in county c. An alternative for the main independent variable is $AvgNatu_c$, the avarage naturalization rate in county c. Both capture the average assimilation effort (X_t in our model). Our

Table A2: Probit estimation: What makes an immigrant fluent in English?

	(1) Without source country variables	(2) With source country variables
PastAvgEnPro	2.570***	2.511***
	(0.0964)	(0.108)
LangDist	-1.181***	-0.583***
	(0.0564)	(0.0686)
NativeIncome	0.0436	0.0564
	(0.0386)	(0.0387)
FemaleDummy	-0.159***	-0.164***
	(0.00846)	(0.00871)
SchoolYear	0.214***	0.216***
	(0.00288)	(0.00292)
AgeImmig	-0.0338***	-0.0337***
	(0.000661)	(0.000698)
YearImmig	0.00930***	0.00848***
	(0.000837)	(0.000787)
MADummy	0.145	0.195
	(0.201)	(0.216)
RelativeGDP		0.332***
		(0.0667)
RelativePop		-0.00724
		(0.0167)
LangShare		1.312***
		(0.0860)
Constant	-2.865***	-3.544***
	(0.404)	(0.422)
N of obs.	172303	163558
pseudo R^2	0.344	0.346

Notes: Robust standard errors clustered by county in parentheses.
$* : p < 0.1,\ ** : p < 0.05,\ *** : p < 0.01$.

Table A3: Marginal effects on the probability that $Enpro_{i,c} = 1$: What matters for the assimilation of immigrants?

	(1) Without source country variables	(2) With source country variables
PastAvgEnPro	0.484***	0.511***
	(0.0175)	(0.0201)
LangDist	-0.222***	-0.119***
	(0.0110)	(0.0142)

Notes: Robust standard errors clustered by county in parentheses.
$* : p < 0.1,\ ** : p < 0.05,\ *** : p < 0.01.$
Evaluated at FemaleDummy=0 (male), MADummy=1 (in MA), and at the mean for other variables.

model predicts an inverted $U-$shaped relationship between $InterMarriage_{i,c}$ and any of these two variables.

A.2.2 Results

We estimate the following model:

$$InterMarriage_{i,c} = \beta_0 + \beta_1 AvgEnPro_c + \beta_1 AvgEnPro_c^2 + Controls + \epsilon_{i,c},$$

and

$$InterMarriage_{i,c} = \beta_0 + \beta_1 AvgNatu_c + \beta_1 AvgNatu_c^2 + Controls + \epsilon_{i,c},$$

where the subscript c denotes county c. Moreover, $\epsilon_{i,c}$ is the error term and $InterMarriage_{i,c}$ is a dummy variable such that:

$$InterMarriage_{i,c} = \begin{cases} 1 & \text{if a native person marriages interracially} \\ 0 & \text{otherwise} \end{cases}$$

Table A4 provides the probit estimation results. Table A5 gives the marginal effects evaluated at different values of $AvgEnPro_c$ or $AvgNatu_c$.[A1] From Table A5, we see that the marginal effects change from positive to negative values as $AvgEnPro_c$ or $AvgNatu_c$ becomes larger. The threshold value (above which the variable becomes negative) is around 0.6-0.7 for $AvgEnPro_c$ and 0.5-0.6 for $AvgNatu_c$. Finally, Figures A1 and A2 plot the marginal effects of the regression in which the

[A1] Note that the minimum value of $AvgEnPro_c$ is 0.44, and its maximum value is 1, whereas the minimum value of $AvgNatu_c$ is 0.14 and its maximum value is 1. Hence, we provide the marginal effect evaluated at the relevant values of these variables in Table A5.

dependent variable is the majority's acceptance rate (measured by the intermarriage rate) and the independent variable of the minority's assimilation is either measured by their English proficiency (Figure A1) or their naturalization rate (Figure A2). Again, we see a clear non-linear relationship. All these empirical results are consistent with the inverted $U-$shaped relationship between $\overline{X}$ and E^* in our model.

Table A4: Probit estimation: Interracial marriage and assimilation of immigrants

	(1) English proficiency	(2) Naturalization
AvgEnPro	4.073**	
	(1.932)	
AvgEnPro^2	-3.277***	
	(1.207)	
AvgNatu		6.943***
		(1.439)
AvgNatu^2		-6.615***
		(1.352)
FemaleDummy	-0.198***	-0.196***
	(0.0148)	(0.0147)
Age	-0.0132***	-0.0133***
	(0.000493)	(0.000492)
SchoolYear	0.0102***	0.00765**
	(0.00372)	(0.00355)
MADummy	0.174	0.207*
	(0.120)	(0.115)
Constant	-2.358***	-3.007***
	(0.764)	(0.404)
N of obs.	368881	368881
pseudo R^2	0.023	0.021

Notes: Robust standard errors clustered by county in parentheses.
$*: p < 0.1$, $**: p < 0.05$, $***: p < 0.01$.

A.3 Evidence from International Social Survey Program

To buttress the previous results, we further use data from the International Social Survey Program (http://w.issp.org/menu-top/home/), denoted as ISSP hereafter.

Table A5: Marginal effects on the probability of interracial marriage: Interracial marriage and the assimilation of immigrants

	(1) English proficiency		(2) Naturalization
AvgEnPro	0.175^{**}	AvgNatu	
=0.4	(0.0696)	=0.1	0.0363^{*}
			(0.0204)
=0.5	0.113	=0.2	0.100^{***}
	(0.0896)		(0.0292)
=0.6	0.0212	=0.3	0.153^{***}
	(0.0794)		(0.0388)
=0.7	-0.0761	=0.4	0.131^{***}
	(0.0541)		(0.0392)
=0.8	-0.153^{***}	=0.5	0.0309^{*}
	(0.0388)		(0.0177)
=0.9	-0.190^{***}	=0.6	-0.0886^{***}
	(0.0375)		(0.0287)
		=0.7	-0.153^{***}
			(0.0417)
		=0.8	-0.132^{***}
			(0.0335)
		=0.9	-0.0652^{***}
			(0.0250)

Notes: Robust standard errors clustered by county in parentheses.
$*: p < 0.1,\ **: p < 0.05,\ ***: p < 0.01$.
For both cases, evaluated at FemaleDummy=0 (male), MADummy=1 (in MA) and at the mean for other variables.

Figure A1: Marginal effects on interracial marriage with respect to English proficiency (with 95 percent confidence intervals)

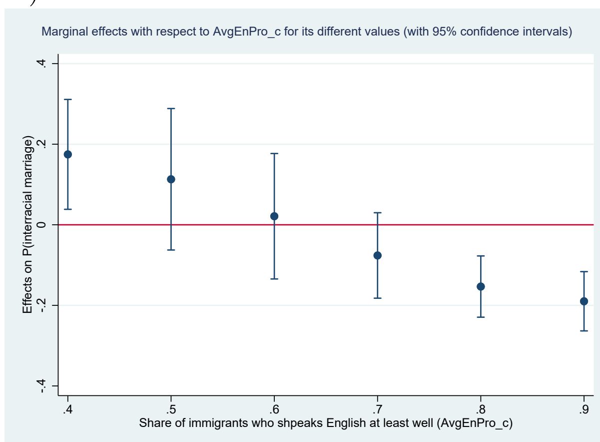
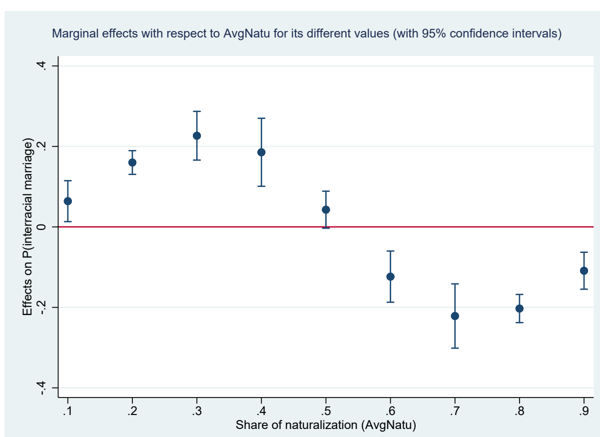


Figure A2: Marginal effects on interracial marriage with respect to naturalization (with 95 percent confidence intervals)



A.3.1 Data

ISSP is an international survey implemented across 31 countries in 1995, 2003, and 2013. We know the citizenship of each person. Thus, we consider people to belong to the majority group if they hold citizenship in the country of their residence. We consider individuals to belong to a minority group if they are not citizens in their country of residence. We specially focus on question v8: "How important is it for you (minority individual) to be able to speak the main language of the country where you live?" and question v40: "How important is it for you (majority individual) that minorities preserve their traditions?" We use the answers to the first question to proxy for the assimilation effort of the immigrant and the answers to the second question to measure the majority's attitude toward the acceptance of immigrants. We create a dummy variable from the answers to these questions. It takes a value of one if their answers represent relatively high effort or high acceptance. See Table A6 for a detailed definition and statistical summary of each variable we use in this dataset.

A.3.2 Empirical results: Minority attitude

We estimate the following model:

$$v8\_d\_min_{i,c} = \beta_0 + \beta_1 PAST\_AVG\_v8_c + Controls + \epsilon_{i,c},$$

where the subscript i, c denotes individual i in country c, $v8\_d\_min_{i,c}$ is the language importance dummy, $PAST\_AVG\_v8_c$ is the past average importance of speaking the national language in country c, and $\epsilon_{i,c}$ is the error term.

Table A7 displays the probit estimation results. It shows that the past average minorities' attitudes toward the national language in the same country is positively correlated with the current attitude.[A2]

A.3.3 Empirical results: Majority's attitude

We estimate the following model:

$$v40\_d\_ma_{i,c} = \beta_0 + \beta_1 AVG\_v8_c + \beta_1 AVG\_v8_c^2 + Controls + \epsilon_{i,c},$$

where the subscript i, c denotes individual i in country c, $v40\_d\_ma_{i,c}$ is the dummy for the importance of accepting the minority's traditions for the majority group, AVG_v8_c is the average

[A2]We use minorities' attitudes in 2003 and 2013 for the dependent variables to show their current attitudes. If the dependent variable is in 2003 (or 2013), the attitude in 1995 and 2003 is used for the past attitude.

Table A6: Descriptive summary statistics of the ISSP dataset

Variables	Description	Mean	S.D.
v8_d_min	Language importance dummy (=1 if a minority's answer to question v8 is "very important/ important")	0.859	-
v40_d_ma	Importance of minorities' culture dummy (=1 if a majority's answer to question v40 is "agree strongly/ agree")	0.377	-
AVG_v8	Average of v8_d_min in the country a person lives in the same period	0.8616	0.181
AVG_v40	Average of v8_d_min in the country a person lives in the same period	0.376	0.222
Male_Dummy	Gender dummy (=1 if a person is male)	0.470	-
MADummy	Metropolitan Area (MA) dummy (=1 if a person locates in a MA)	0.362	-
AGE	Age	46.290	17.675
HiEduDummy	High education dummy	0.201	-
LowEduDummy	Low education dummy	0.358	-
MarryDummy	Marriage dummy (=1 if a person is married)	0.579	-
KidsDummy	Kids dummy (=1 if a person has kid(s))	0.413	-
WorkDummy	=1 if a person works	0.577	-
GNI	Gross national income of the country a person lives in the same period (US dollars, in natural logarithm)	10.144	0.833
y03Dummy	Dummy for 2003 samples (=1 if the sample is observed in 2003)	0.410	-
y13Dummy	Dummy for 2013 samples (=1 if the sample is observed in 2013)	0.335	-

Notes: Details of the key questions are as follows.

v8: How important is it to be able to speak the language of the country in which a person lives?

v40: Help minorities to preserve traditions.

A person who (does not) has citizenship is identified as a majority (minority).

The GNI of each country is provided by the World Bank, which is considered as the majority's average income.

All other data come from the ISSP dataset.

Table A7: Probit estimations: Minority attitudes

PAST_AVG_v8	3.404883***
	(0.2349089)
GNI	0.3653194***
	(0.0792193)
MaleDummy	-0.0145108
	(0.093021)
MADummy	-0.146976
	(0.0976096)
AGE	0.0004719
	(0.0033999)
HiEduDummy	-0.1633532
	(0.1098569)
LowEduDummy	-0.0848859
	(0.1202703)
MarryDummy	-0.0301673
	(0.1003935)
WorkDummy	-0.1509855
	(0.1008534)
KidsDummy	0.0234704
	(0.102949)
y03Dummy	0.3267093***
	(0.1231838)
constant	-5.304951***
	(0.8543046)
N	1542
R^2	0.2684

Notes: Robust standard errors clustered by country in parentheses.
$*$: $p < 0.1$, $**$: $p < 0.05$, $***$: $p < 0.01$
v8_d of each minority is estimated by PAST_AVG_v8,
which is the country average of v8_d among minorities in the previous period.
Therefore, only the samples in 2003 and 2013 are used for estimation.
The data in 1995 are only used for making PAST_AVG_v8.

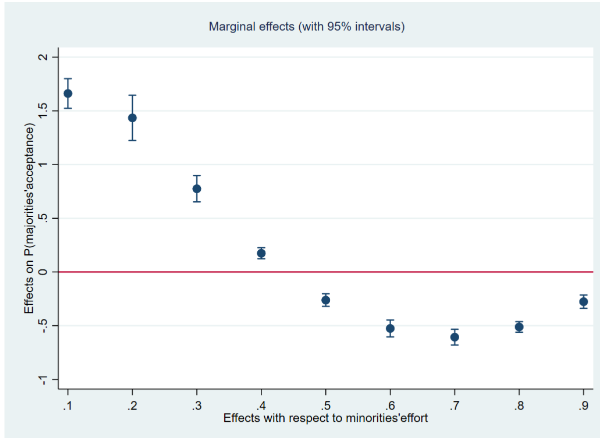
language importance dummy for non-citizens, and $\epsilon_{i,c}$ is the error term. Table A8 displays the probit estimation results. We again observe an inverted $U-$shaped pattern between AVG_v8_c and $v40\_d\_ma_{i,c}$.

Table A8: Probit estimations: Majority's attitude

AVG_v8	1.488366 ***
	(0.1702186)
AVG_v8^2	-1.628524***
	(0.1355203)
Male_Dummy	-0.0148283
	(0.0122554)
MADummy	0.0579953***
	(0.0136448)
Age	-0.0031286***
	(0.0004082)
HiEduDummy	0.1651169***
	(0.0156852)
LowEduDummy	0.0187982
	(0.0143362)
MarryDummy	-0.0549825***
	(0.013383)
WorkDummy	-0.131982***
	(0.0134133)
KidDummy	0.0870884***
	(0.0138308)
y03Dummy	0.2598087***
	(0.0174454)
y13Dummy	0.1241997***
	(0.0141266)
const.	-0.6473142***
	(0.0572802)
N of obs.	48389
pseudo R^2	0.0218

Notes: Robust standard errors in parentheses
$* : p < 0.1,\ ** : p < 0.05,\ *** : p < 0.01.$

Figure A3: Marginal effects of minorities' efforts on majorities' acceptance



B Proofs of the results in the main text

Proof of Proposition 1

The first-order condition for each i is given by:

$$\begin{aligned} 0 &= \frac{\partial\, U_i|_{\text{Y given}}}{\partial x_i} \qquad \text{(B.1)} \\ &= \alpha_i y'(x_i) - \rho\left[p'(x_i)(\overline{x} - x_i) - p(x_i)\right] + \sigma p'(x_i)(Z - Y) - h(X_0)c'(x_i). \end{aligned}$$

The second-order condition is given by

$$\begin{aligned} 0 &> \frac{\partial^2\, U_i|_{\text{Y given}}}{\partial x_i^2} \qquad \text{(B.2)} \\ &= \alpha_i y''(x_i) - \rho\left[p''(x_i)(\overline{x} - x_i) - 2p'(x_i)\right] + \sigma p''(x_i)(Z - Y) - h(X_0)c''(x_i), \end{aligned}$$

which we assume to hold throughout the paper.[A3]

Equilibrium in this baseline framework is given by a tupple $(x_1, x_2, ..., x_m, Y)$ that satisfies the first order condition (B.1) for all i and $Y = \left(\sum_{i=1}^m y(x_i)\right)/m$.

From the definition of the average income, we know that

$$\frac{\partial Y}{\partial \alpha_i} = \frac{1}{m}\sum_{j=1}^m y'(x_j)\frac{\partial x_j}{\partial \alpha_i}. \qquad \text{(B.3)}$$

We incorporate (B.3) in totally differentiating the first order conditions (B.1) for individuals i and j to obtain

$$\begin{aligned} \frac{\partial x_i}{\partial \alpha_i} &= \frac{\sigma p'(x_i)\frac{\partial Y}{\partial \alpha_i} - y'(x_i)}{\frac{\partial^2\, U_i|_{\text{Y given}}}{\partial x_i^2}}, \qquad \text{(B.4)} \\ \frac{\partial x_j}{\partial \alpha_i} &= \frac{\sigma p'(x_j)\frac{\partial Y}{\partial \alpha_i}}{\frac{\partial^2\, U_j|_{\text{Y given}}}{\partial x_j^2}}. \end{aligned}$$

Plugging these into (B.3), we can see that

$$\frac{\partial Y}{\partial \alpha_i} = -\frac{\left(y'(x_i)\right)^2}{m\frac{\partial^2\, U_i|_{\text{Y given}}}{\partial x_i^2}} + \frac{\sigma}{m}\frac{\partial Y}{\partial \alpha_i}\sum_{j=1}^m \frac{p'(x_j)y'(x_j)}{\frac{\partial^2\, U_j|_{\text{Y given}}}{\partial x_j^2}},$$

[A3] This is satisfied if $p''(x_i) \leq 0$ and ρ is sufficiently small.

which yields

$$\frac{\partial Y}{\partial \alpha_i} = -\frac{\frac{(y'(x_i))^2}{\partial^2 U_i|_{\text{Y given}}/\partial x_i^2}}{m - \sigma \sum_{j=1}^m \frac{p'(x_j)y'(x_j)}{\partial^2 U_j|_{\text{Y given}}/\partial x_j^2}} > 0, \tag{B.5}$$

where the inequality comes from the second order condition $\partial^2 U_i|_{\text{Y given}} / \partial x_i^2 < 0$. From this and (B.4), we readily know that

$$\frac{\partial x_j}{\partial \alpha_i} < 0. \tag{B.6}$$

Equations (B.4) and (B.5) yield

$$\begin{aligned}
\frac{\partial x_i}{\partial \alpha_i} &= -\frac{1}{\partial^2 U_i|_{\text{Y given}} / \partial x_i^2} \left(\frac{\sigma p'(x_i) \frac{(y'(x_i))^2}{\partial^2 U_i|_{\text{Y given}}/\partial x_i^2}}{m - \sigma \sum_{j=1}^m \frac{y'(x_j)p'(x_j)}{\partial^2 U_j|_{\text{Y given}}/\partial x_j^2}} + y'(x_i) \right) \\
&= -\frac{1}{\partial^2 U_i|_{\text{Y given}} / \partial x_i^2} \frac{y'(x_i)}{m - \sigma \sum_{j=1}^m \frac{p'(x_j)y'(x_j)}{\partial^2 U_j|_{\text{Y given}}/\partial x_j^2}} \\
&\times \left(\sigma \frac{p'(x_i)y'(x_i)}{\partial^2 U_i|_{\text{Y given}} / \partial x_i^2} + m - \sigma \sum_{j=1}^m \frac{p'(x_j)y'(x_j)}{\partial^2 U_j|_{\text{Y given}} / \partial x_j^2} \right) \\
&= -\frac{1}{\partial^2 U_i|_{\text{Y given}} / \partial x_i^2} \frac{y'(x_i)}{m - \sigma \sum_{j=1}^m \frac{p'(x_j)y'(x_j)}{\partial^2 U_j|_{\text{Y given}}/\partial x_j^2}} \left(m - \sigma \sum_{j \neq i} \frac{p'(x_j)y'(x_j)}{\partial^2 U_j|_{\text{Y given}} / \partial x_j^2} \right) > 0.
\end{aligned} \tag{B.7}$$

Equations (B.5), (B.6), and (B.7) prove part (*i*) of the proposition.

To prove parts (*ii*) and (*iii*) of the proposition, we can obtain the effects of other parameters as follows. Consider a change in parameter, ξ ($= X_0$, ρ, σ, or Z). From the definition of the average income, we know that

$$\frac{\partial Y}{\partial \xi} = \frac{1}{m} \sum_{j=1}^m y'(x_j) \frac{\partial x_j}{\partial \xi}. \tag{B.8}$$

We incorporate (B.8) in totally differentiating the first order condition (B.1) to obtain

$$\frac{\partial x_i}{\partial \xi} = \frac{\sigma p'(x_i) \frac{\partial Y}{\partial \xi} - \frac{\partial^2 U_i|_{\text{Y given}}}{\partial x_i \partial \xi}}{\frac{\partial^2 U_i|_{\text{Y given}}}{\partial x_i^2}}. \tag{B.9}$$

Plugging this into (B.8) and rearranging it yield

$$\frac{\partial Y}{\partial \xi} = -\frac{\sum_{j=1}^m \frac{y'(x_j)}{\partial^2 U_j|_{\text{Y given}}/\partial x_j^2} \frac{\partial^2 U_j|_{\text{Y given}}}{\partial x_j \partial \xi}}{m - \sigma \sum_{j=1}^m \frac{p'(x_j)y'(x_j)}{\partial^2 U_j|_{\text{Y given}}/\partial x_j^2}}. \tag{B.10}$$

Once we obtain $\partial^2 \left. U_i \right|_{\text{Y given}} / \partial x_i \partial \xi$, (B.10) determines $\partial Y / \partial \xi$, which, combined with (B.9), determines $\partial Y / \partial \xi$. We can see that

$$\frac{\partial^2 \left. U_i \right|_{\text{Y given}}}{\partial x_i \partial X_0} = -h'(X_0) c'(x_i) \geq 0, \tag{B.11}$$
$$\frac{\partial^2 \left. U_i \right|_{\text{Y given}}}{\partial x_i \partial \rho} = -p'(x_i)(\overline{x} - x_i) + p(x_i),$$
$$\frac{\partial^2 \left. U_i \right|_{\text{Y given}}}{\partial x_i \partial \sigma} = p'(x_i)(Z - Y) \geq 0,$$
$$\frac{\partial^2 \left. U_i \right|_{\text{Y given}}}{\partial x_i \partial Z} = \sigma p'(x_i) > 0.$$

After arranging equations (B.9), (B.10), and (B.11), we can observe that the signs of $\partial Y / \partial \rho$ and $\partial x_i / \partial \rho$ are ambiguous, while $\partial Y / \partial \sigma \geq 0$ and $\partial x_i / \partial \sigma \geq 0$, and $\partial Y / \partial Z > 0$ and $\partial x_i / \partial Z > 0$. Moreover, we can see that $\partial Y / \partial X_0 \geq 0$. Therefore, although $\partial x_i / \partial X_0$ can be positive or negative, $\partial Y / \partial X_0 \geq 0$ implies that $\partial x_i / \partial X_0 \geq 0$ must hold true for some i. The above arguments prove parts (*ii*) and (*iii*) of the proposition. ▮

Proof of Proposition 2

Plugging (B.1) into the derivative of (3) with respect to x_i yields

$$\left. \frac{\partial W}{\partial x_i} \right|_{\text{equilibrium}} = -\frac{y'(x_i)}{m} \sum_{j=1}^{m} p(x_j) \leq 0,$$

where the equality holds true if and only if $x_i = \overline{x}$ for all i.[A4] Put differently, unless the first order conditions yields the corner solution $\overline{x}$ i.e., the maximum possible value of x_i, the equilibrium assimilation effort is inefficiently high. ▮

Proof of Proposition 3

Before proving Proposition 3, let us state and prove the following lemmas:

Lemma B1 *Suppose that $\alpha_i = \alpha$ for all i. Assume also that*

$$\frac{\gamma}{1+\delta} > \frac{2\rho}{\delta}, \tag{B.12}$$

[A4]Note that the envelop theorem holds because $\partial U_i / \partial x_i = 0$ according to (B.1), and hence only the externalities remain.

and

$$\Theta \equiv \alpha\theta - \rho + \frac{\sigma}{\delta}(Z - y) > 0. \tag{B.13}$$

Then, $dX_t/dX_{t-1} > 0$ *and* $d^2X_t/dX_{t-1}^2 > 0$ *if and only if* $2\rho > \sigma\theta$.

Proof of Lemma B1

Before proving Lemma B1, let us first determine the optimization problem of assimilation effort in the general case and then in the symmetric case.

General case when agents have different αs

Under the specifications of the different functions, the first order condition (B.1) becomes

$$x_{it} = \frac{\alpha_i\theta - \rho + \frac{\sigma}{\delta}(Z - Y_t)}{\frac{\gamma}{1+X_{t-1}} - \frac{2\rho}{\delta}}. \tag{B.14}$$

The second order condition now is given by $\gamma/(1 + X_{t-1}) > 2\rho/\delta$. Because $0 \leq X_{t-1} \leq \delta$, this holds true if (B.12) holds. Including the case of a corner solution, x_{it} is determined by

$$x_{it} = \min\left[\delta, \max\left[0, \frac{\alpha_i\theta - \rho + \frac{\sigma}{\delta}(Z - Y_t)}{\frac{\gamma}{1+X_{t-1}} - \frac{2\rho}{\delta}}\right]\right].$$

Let us now focus on the symmetric case.

Symmetric case

Assume that $\alpha_i = \alpha, \forall i$. In this case, an interior solution is given by (B.14) with $\alpha_i = \alpha$ for all i. By summing it up and solving it with respect to $X_t \equiv (\sum_{j=1}^{m} x_{jt})/m$, we obtain

$$X_t = \frac{\alpha\theta - \rho + \frac{\sigma}{\delta}(Z - y)}{\frac{\gamma}{1+X_{t-1}} + \frac{\sigma\theta - 2\rho}{\delta}}. \tag{B.15}$$

Equation (B.15) represents the law of motion of the average assimilation effort, X_t. Because the denominator of (B.15) is positive by assumption (B.12), X_t determined by (B.15) is positive if condition (B.13) holds. Because Θ is the average value of the marginal returns of assimilation evaluated at $x_i = 0$, a negative value of Θ implies that minority individuals, on average, have no incentive to assimilate. In order to avoid such an extreme situation, we assume (B.13) throughout the rest of the paper. To close the model, we impose the steady state condition

$$X_t = X_{t-1}. \tag{B.16}$$

Let us now prove Lemma B1. Differentiating (B.15) with respect to X_t yields

$$\frac{dX_t}{dX_{t-1}} = \frac{\gamma\Theta}{\left[\gamma + \frac{\sigma\theta - 2\rho}{\delta}(1 + X_{t-1})\right]^2} > 0.$$

By differentiating this with respect to X_t, we obtain

$$\frac{d^2X_t}{dX_{t-1}^2} = \frac{2\gamma\Theta}{\left[\gamma + \frac{\sigma\theta - 2\rho}{\delta}(1 + X_{t-1})\right]^3} \frac{2\rho - \sigma\theta}{\delta},$$

which proves the lemma. ∎

Condition (B.12) guarantees that there is a unique solution for the optimization problem of agent i (that is, the second-order condition is satisfied) while (B.13) ensures that X_t is always strictly positive.

In this dynamic framework, a steady state equilibrium is given by a tuple (x_i^*, X^*, Y^*). These equilibrium variables are determined by the first order condition (B.14), the law of motion of X_t (B.15), and $Y_t = \left(\sum_{i=1}^m y(x_i)\right)/m$, combined with the steady state condition (B.16). In the case of a corner solution, the law of motion of X_t (B.15) is replaced by $X_t = 0$ or $X_t = 1$. Moreover, we impose stability, which requires that $dX_t/dX_{t-1}|_{X_t = X_{t-1}} < 1$ in the case of an interior solution, that X_t determined by (B.15) for $X_{t-1} = 1$ is larger than δ in the case of a corner solution with $X_t = 1$, and that X_t determined by (B.15) for $X_{t-1} = 0$ is smaller than 0 in the case of a corner solution with $X_t = 0$. Under assumptions (B.12) and (B.13), X_t determined by (B.15) for $X_{t-1} = 0$ becomes positive, which implies that we have no stable equilibrium with a corner solution with $X_t = 0$.

Now let us prove Proposition 3:

Suppose (B.12), (B.13), and $2\rho > \sigma\theta$ hold true. Then, from Figure 1, we know that multiple stable steady state equilibria exist if and only if:

(i) X_t is determined by (B.15) and X_t at $X_{t-1} = 0$ as well as X_t at $X_{t-1} = \delta$ are both positive, that is, $X_t|_{X_{t-1}=0} > 0$ and $X_t|_{X_{t-1}=\delta} > 0$,

(ii) the simultaneous equations (B.15) and (B.16) have two distinct real roots,

(iii) the two distinct real roots are within the interval $[0, \delta]$.

Note first that (i) hold true because

$$X_t|_{X_{t-1}=0} = \frac{\Theta}{\gamma + \frac{\sigma\theta - 2\rho}{\delta}} > 0,$$

$$X_t|_{X_{t-1}=\delta} = \frac{\Theta}{\frac{\gamma}{1+\delta} + \frac{\sigma\theta - 2\rho}{\delta}} > 0,$$

where these inequalities use the assumptions (B.12) and (B.13). Moreover, if we assume (ii), that is, the existence of interior solutions (two distinct real roots) for (B.15) and (B.16), they are given by

$$X^* = \frac{-\left(\Theta - \gamma + \frac{2\rho-\sigma\theta}{\delta}\right) \pm \sqrt{\left(\Theta - \gamma + \frac{2\rho-\sigma\theta}{\delta}\right)^2 - \frac{4\Theta(2\rho-\sigma\theta)}{\delta}}}{\frac{2(2\rho-\sigma\theta)}{\delta}}. \tag{B.17}$$

Therefore, $X^* \geq 0$ holds true if and only if

$$-\left(\Theta - \gamma + \frac{2\rho - \sigma\theta}{\delta}\right) > 0 \Leftrightarrow \Theta < \gamma - \frac{2\rho - \sigma\theta}{\delta}. \tag{B.18}$$

$X^* \leq \delta$ holds true if and only if

$$-\left(\Theta - \gamma + \frac{2\rho - \sigma\theta}{\delta}\right) \pm \sqrt{\left(\Theta - \gamma + \frac{2\rho - \sigma\theta}{\delta}\right)^2 - \frac{4\Theta(2\rho - \sigma\theta)}{\delta}} \leq 2(2\rho - \sigma\theta).$$

Under (B.18), the left hand side of the above inequality is smaller than $-2\left(\Theta - \gamma + \frac{2\rho-\sigma\theta}{\delta}\right)$, which is equal or smaller than $2(2\rho - \sigma\theta)$ if and only if

$$\Theta \geq \gamma - \left(1 + \frac{1}{\delta}\right)(2\rho - \sigma\theta). \tag{B.19}$$

Hence, if we assume (ii), (B.18) and (B.19) ensure (iii) to hold true.

Finally, we provide one set of sufficient conditions for (ii) to hold true under (B.18) and (B.19). From (B.17), X^* takes two distinct real values if and only if

$$\left(\Theta - \gamma + \frac{2\rho - \sigma\theta}{\delta}\right)^2 > \frac{4\Theta(2\rho - \sigma\theta)}{\delta}. \tag{B.20}$$

If we define $\Gamma(\Theta)$ as

$$\Gamma(\Theta) = \frac{\left(\Theta - \gamma + \frac{2\rho-\sigma\theta}{\delta}\right)^2}{\Theta},$$

equation (B.20) is satisfied if and only if $\Gamma(\Theta) > \frac{4(2\rho-\sigma\theta)}{\delta}$. Under (B.18) and (B.19), we can see that $\Gamma'(\Theta) < 0$. Consider ω that satisfies $0 < \omega < \delta$. Then, there exists an interval $[\gamma - \left(1 + \frac{1}{\delta}\right)(2\rho - \sigma\theta), \gamma - \left(\frac{\omega}{\delta} + \frac{1}{\delta}\right)(2\rho - \sigma\theta)]$. Suppose that Θ satisfies

$$\gamma - \left(1 + \frac{1}{\delta}\right)(2\rho - \sigma\theta) \leq \Theta \leq \gamma - \left(\frac{\omega}{\delta} + \frac{1}{\delta}\right)(2\rho - \sigma\theta). \tag{B.21}$$

Then, $\Gamma(\Theta) > \frac{4(4\rho-\sigma\theta)}{\delta}$ holds true if

$$\Gamma\left(\gamma - \left(\frac{\omega}{\delta} + \frac{1}{\delta}\right)(2\rho - \sigma\theta)\right) \geq \frac{4(2\rho - \sigma\theta)}{\delta},$$

which is written as

$$\left(1 + \omega + \frac{\omega^2}{4}\right)\frac{(2\rho - \sigma\theta)}{\delta} \geq \gamma. \tag{B.22}$$

We thus know that there exist multiple stable steady state equilibria if (B.21), and (B.22) hold true for some ω and Θ [A5], which proves Proposition 3. ∎

Proof of Proposition 4

From Figure 1, we can see that the interior solution of X^* becomes larger as the locus of (B.15) moves upwards. Because (B.15) implies that $\partial X_t/\partial(Z-y) > 0$, $\partial X_t/\partial\alpha > 0$, $\partial X_t/\partial\gamma < 0$, we know that $\partial X^*/\partial(Z-y) > 0$, $\partial X^*/\partial\alpha > 0$, $\partial X^*/\partial\gamma < 0$. Moreover, because the signs of $\partial X_t/\partial\theta$, $\partial X_t/\partial\rho$, $\partial X_t/\partial\sigma$, and $\partial X_t/\partial\delta$ are not determined, the effects of these parameters on X^* are ambiguous. ∎

Proof of Proposition 5

The welfare criterion is given by the sum of welfare (3) over time:

$$TW = \sum_{t=0}^{\infty} \beta^t W_t, \tag{B.23}$$

where β is the discount factor satisfying that $0 < \beta < 1$. Plugging (B.15) into the derivative of (B.23) with respect to x_{it} yields

$$\begin{aligned} \left.\frac{\partial TW}{\partial x_{it}}\right|_{\text{equilibrium}} &= -\beta^t \frac{\partial Y_t}{\partial x_{it}} \sum_{j=1}^{m} p(x_{jt}) - \beta^{t+1}\frac{\partial X_t}{\partial x_{it}} h'(X_t) \sum_{j=1}^{m} c(x_{jt+1}) \\ &= -\beta^t \frac{\sigma\theta}{m} \sum_{j=1}^{m} \left(\frac{x_{jt}}{\delta}\right) + \beta^{t+1} \frac{\gamma}{m(1+X_t)^2} \sum_{j=1}^{m} \frac{x_{jt+1}^2}{2}. \end{aligned} \tag{B.24}$$

The first term of the right hand side of (B.24) shows the negative externality of assimilation to other individuals in the same generation. However, a large assimilation effort decreases assimilation cost of next generation and then it is beneficial for them. This positive externality is described by

[A5] The conditions given by (B.18) and (B.19) are included in (B.21) the second term.[A6] Therefore, the effect of assimilation is negative in the short run while positive in the long-run, and the sign of the aggregated externalities depends on value of β. ▮

Proof of Proposition 6

By differentiating (6), we obtain:

$$\frac{\partial e_k^*}{\partial X} = \frac{q'(X)(\overline{x} - X) - q(X)}{d''(e_k)}.$$

Evaluating this at $X = 0$ or at $X = \overline{x}$, we obtain

$$\left.\frac{\partial e_k^*}{\partial X}\right|_{X=0} = \frac{q'(0)\overline{x}}{d''(e_k)} > 0 \qquad \text{and} \qquad \left.\frac{\partial e_k^*}{\partial X}\right|_{X=\overline{x}} = -\frac{q(\overline{x})}{d''(e_k)} < 0,$$

which proves the proposition. ▮

Proof of Proposition 7

The welfare criterion is given by

$$TW_a = \sum_{t=0}^{\infty} \beta^t W_{at}, \tag{B.25}$$

where W_{at} incorporates both minority's and majority's utility function. It is defined as

$$W_{at} = \sum_{j=1}^{m} U_{jt} + \sum_{k=1}^{n} V_{kt}.$$

Plugging (7) into the derivatives of (B.25) with respect to x_{it} and e_{it} yields

$$\begin{aligned}
\left.\frac{\partial TW_a}{\partial x_{it}}\right|_{\text{equilibrium}} &= \beta^t \left[-\frac{\sigma\theta}{m} \sum_{j=1}^{m} \frac{x_{jt}}{\delta} + \frac{\beta\gamma(\lambda - \sum_{k=1}^{n} e_{kt}/n)}{m(1+X_t)^2} \sum_{j=1}^{m} \frac{x_{jt+1}^2}{2} \right] \\
&\quad - \beta^t \sum_{k=1}^{n} \varepsilon(\lambda - e_{kt}) \frac{\delta - 2X_t}{m}, \\
\left.\frac{\partial TW_a}{\partial e_{it}}\right|_{\text{equilibrium}} &= \beta^t \frac{\gamma}{2n(1+X_{t-1})} \sum_{j=1}^{m} x_{jt}^2 > 0.
\end{aligned}$$

[A6] We assume that only individual i changes x_{it} in generation t, while behaviors of other individuals are fixed. However, change in x_{it} may affect the behaviors of individuals of the current and future generations.

If we focus on the symmetric case where $\alpha_i = \alpha$, $\forall i$, this yields

$$\left.\frac{\partial TW_a}{\partial x_{it}}\right|_{\text{equilibrium}} = \beta^t \left[-\sigma\theta \frac{X_t}{\delta} + \frac{\beta(\lambda - E_t)\gamma X_{t+1}^2}{2(1+X_t)^2} - \varepsilon(\lambda - E_t)(\delta - 2X_t)\frac{n}{m} \right].$$

In this equation, the third term in the brackets is negative and decreasing in n/m. Hence, the results from the proposition follow. ∎

Proof of Proposition 8

Assume that the government controls $\overline{E}$. Then, the assimilation effort is endogenously determined by (9). Because we impose stability, we know that $\partial X^*(X^*, \overline{E})/\partial X^* < 1$. Therefore, $dX^*/d\overline{E} = \frac{\partial X^*(X^*,\overline{E})/\partial \overline{E}}{1 - \partial X^*(X^*,\overline{E})/\partial X^*} > 0$ holds true. Assume next that the government controls $\overline{X}$. Then, the degree of acceptance is endogenously determined by (10). Then, we obtain $dE^*/d\overline{X}$. ∎

Proof of Proposition 9

The effects of the planner's policy on the welfare evaluated at an interior equilibrium (i.e., at $\overline{E} = E^*$ and $\overline{X} = X^*$) are described as follows:

$$\left.\frac{\partial TW_a}{\partial \overline{X}}\right|_{\text{equilibrium}} = \underbrace{-\frac{m\sigma\theta X^*}{\delta} + \frac{dE^*}{d\overline{X}}\frac{\gamma m X^{*2}}{2(1+X^*)}}_{\text{current minority's welfare } (<0 \text{ if } X^*>\delta/2, \text{ ambiguous otherwise})}$$

$$+\frac{\beta}{1-\beta}\left[\underbrace{-\frac{m\sigma\theta X^*}{\delta} + \frac{dE^*}{d\overline{X}}\frac{\gamma m X^{*2}}{2(1+X^*)} + \frac{\gamma(\lambda - E^*)X^{*2}}{2(1+X^*)^2}}_{\text{future minority's welfare (ambiguous)}}\right]$$

$$\underbrace{-\frac{n\varepsilon(\lambda - E^*)}{1-\beta}\frac{\delta - 2X^*}{m}}_{\text{majority's welfare } (>0 \text{ iff } X^*>\delta/2, <0 \text{ otherwise})},$$

$$\left.\frac{\partial TW_a}{\partial \overline{E}}\right|_{\text{equilibrium}} = \underbrace{-\frac{dX^*}{d\overline{E}}\frac{m\sigma\theta X^*}{\delta} + \frac{\gamma m X^{*2}}{2(1+X^*)}}_{\text{current minority's welfare (ambiguous)}}$$

$$+\frac{\beta}{1-\beta}\left[\underbrace{-\frac{dX^*}{d\overline{E}}\frac{m\sigma\theta X^*}{\delta} + \frac{\gamma m X^{*2}}{2(1+X^*)} + \frac{dX^*}{d\overline{E}}\frac{\gamma(\lambda - E^*)X^{*2}}{2(1+X^*)^2}}_{\text{future minority's welfare (ambiguous)}}\right]$$

$$\underbrace{-\frac{dX^*}{d\overline{E}}\frac{n\varepsilon(\lambda - E^*)}{1-\beta}\frac{\delta - 2X^*}{m}}_{\text{majority's welfare } (>0 \text{ iff } X^*>\delta/2, <0 \text{ otherwise})}.$$

In each equation, the total effect is decomposed into the effects on the minority's welfare of the current generation (i.e. the welfare of minority individuals who exist in the period when the policy begins), those on the minority's welfare of the future generation, and those on the majority's welfare.[A7] They have not only the direct effects of the exogenously controlled variable, but also the indirect effects caused by changes in the endogenous variables. Recall Proposition 8 showing that $dX^*/d\bar{E}$ is always positive while $dE^*/d\bar{X}$ is positive (resp. negative) if and only if X^* is larger (resp. smaller) than $\delta/2$. Therefore, the policy effects on the minority's welfare depend on generation, yielding the proposition. ∎

[A7] We consider these effects around the interior stable steady-state equilibrium in which both X and E are determined endogenously. Therefore, the Envelope theorem can be applied for each group's welfare.