Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 175 (2018) 374–414

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Cultural leader and the dynamics of assimilation ☆

Thierry Verdier [a], Yves Zenou [b,*]

[a] PSE and Ecole des Ponts ParisTech, PUC-Rio and CEPR, France
[b] Monash University, IFN and CEPR, Australia

Received 16 July 2016; final version received 26 January 2018; accepted 31 January 2018
Available online 6 February 2018

Abstract

This paper studies the population dynamics of cultural traits in a model of intergenerational cultural transmission with a perfectly-forward looking cultural leader. We show that there exists a threshold size in terms of population above which the cultural leader becomes active. We also show that a policy affecting some key parameters (such as the cost of providing the religious good) has a different impact in the short run and in the long run due to over-reactions or under-reactions of the different cultural groups. Finally, we study the cultural competition between two forward-looking cultural leaders with opposite objectives. We show that the steady-state cultural equilibrium depends on the time preference structure of the two leaders.

JEL classification: J13; J15; Z10

Keywords: Cultural substituability; Integration; Forward-looking leader

1. Introduction

The social integration of immigrants into a host society is a very heated debate, especially in Europe. It is not clear, however, what "integration" means. Immigrants often have a relationship with at least two cultures, their ethnic cultural background (the minority culture) and the majority

☆ We thank the editor, an associate editor as well as two anonymous referees for very helpful comments. This research was undertaken under the project TECTACOM. Financial support from the European Research Council (ERC) grant number 324004 is gratefully acknowledged.

* Corresponding author.
E-mail addresses: thierry.verdier@ens.fr (T. Verdier), yves.zenou@monash.edu (Y. Zenou).

https://doi.org/10.1016/j.jet.2018.01.019
0022-0531/

culture in the country where they live. Different European countries have different views of their integration policy and in later years, there has been a marked change in the attitude to the question of diversity (integration) versus assimilation. Certain countries consider it to be a successful integration policy when immigrants leave their cultural background and are "assimilated" into the new culture. An obvious example is the French law on religious symbols at school. Since the fall of 2004, religious symbols have been forbidden in French schools. This means that Muslim girls are no longer allowed to wear veils at school. But such a law forces immigrants to leave their original culture and choose the French culture. Other countries consider that a successful integration policy is that immigrants can keep their original culture while also accepting the new culture (or at least not rejecting it). This is the British model. Language instruction is an important part of this. By improving their knowledge of the language, immigrants are assisted in becoming more quickly adjusted to the new country without abandoning their own culture.

In this debate on the integration of immigrants in the host country, there is one important issue that has been somewhat neglected by policy discussions: the *role of cultural leaders* and their influence as coordinating socialization agents. It is, indeed, well-documented that cultural leaders or the cultural institutions that they represent (churches, mosques, schools, ethnic associations) have an important impact on the integration and assimilation of immigrants in Europe and the United States. For example, in the United States, sermons, lectures, and discussions held in places of worship on immigration influence the ways that parishioners think about immigration policy (Nteta and Wallsten, 2012). Within Islam, contemporary pronouncements by clerics can have substantial sway among lay Muslims, defining norms of acceptability and permissibility for the entire range of human action (Nielsen, 2012). In Germany, imams paid by the Turkish state transmit their religious knowledge by means of sermons in the mosque, in particular, Friday sermons which offer a one-way flow of knowledge from the imam to the believers (Yurdakul and Yükleyen, 2009). Some religious leaders even give anti-German speeches in mosques. In his Friday sermon at Mevlana Mosque, a well-established mosque in Kreuzberg directed by the Islam Federation in Berlin (IFB), an Imam was filmed by German TV station ZDF giving such an address. Among other things, the imam said Germans will go to hell, they do not shave their armpits and they stink. After the broadcast, the imam apologized, but IFB removed him from his position (Yükleyen and Yurdakul, 2011).

In this paper, we would like to investigate these issues by focusing on the role of cultural leaders in the integration of immigrants in the host country. Indeed, to have an effective integration policy, one needs to understand the way ethnic minorities form their identity, get culturally organized and how they identify themselves with different role models. Surprisingly, in economics, there is little research on the role of cultural leaders in the formation of cultural identity of immigrants, even though community leaders are common in immigrant communities.

For that, we develop a model where individuals get socialized to specific cultural traits in two ways. The first channel is the usual *decentralized* evolutionary mechanism working through parents direct inculturation efforts (*vertical socialization*) and social exposure to peers (*oblique socialization*). The second mechanism involves community leaders or cultural institutions with a longer-term view and more encompassing vision on cultural dynamics at the *group* level, which therefore implies a more *centralized* process of socialization.

To be more precise, individuals can be of two types (a or b), for example, "religious" or "secular", "ethnic" or "mainstream", etc. They enjoy utility from consuming a public good related to their trait and the incentives to choose their preferred action increase with the amount of specific public good provided for that trait. For example, for a religious Muslim person, the public good could be a mosque or an Islamic school, while for members of a specific ethnic group, it would be community infrastructures facilitating social interactions and activities with other co-ethnic individuals. As in Bisin and Verdier (2000), parents choose how much effort to exert to socialize their offsprings by maximizing their children's utility with imperfect empathy. In such a context, the evolution of intergenerational cultural transmission of a specific trait takes the form of a typical replicator dynamics equation that depends on the difference in efforts between the parents with different traits, which, in turn, depends on the group specific public good provided by community leaders.

We embed this evolutionary *decentralized* transmission mechanism at the family and peer level with the existence of community leaders or cultural institutions that act as *centralized* socializing agents. We assume the leaders to be *perfectly forward looking* and to exert their socialization activity in order to maximize a well-defined *intertemporal objective* over certain cultural characteristics of their group. This implies, in particular, that these socializing agents fully internalize the effects of their efforts on the population dynamics of cultural evolution. These features thus capture an important dimension that differentiates community leaders (or the cultural institution they represent) from other socialization patterns associated with families and peers. Typically, cultural leaders have *longer time horizons* and more encompassing visions about the cultural evolution of their group than any other individual member. As a consequence, their socialization efforts are less myopic and more purposeful than the transmission patterns emerging from families and peers. These features clearly matter for the resulting equilibrium cultural trajectory that combines the various channels of socialization (family, peers, leaders), which help sustain a given trait in a population. In this context, our contribution is threefold.

First, by considering the case of one cultural leader only (say from group a), we provide a full characterization of the optimal cultural control trajectory, which depends on the initial conditions of the population. Technically, the centralized socialization problem of a leader is closely related to an epidemiology optimal control problem in which an intertemporal objective function is maximized under a replicator dynamic equation characterizing the pattern of cultural diffusion inside the population. This problem naturally involves dynamic non-convexities and is difficult to solve by the traditional first-order approach. With a linear bang-bang socialization technology for the leader and using a Most Rapid Path Approach (MRAP) methodology developed by Spence and Starrett (1975), we are, however, able to fully characterize the cultural dynamics of the population, not only at the steady state, but also along the whole transition path.

Second, our analysis allows for some interesting comparative dynamics along the transition path of cultural evolution. Specifically, we investigate the impact of various shocks on the dynamics of population, contrasting the short-run effects from the long-run impacts of these shocks on the cultural trajectory.

Our last contribution consists in studying the *cultural competition* between two forward-looking cultural leaders (for example, a religious and a secular leader who have opposite views on the integration and assimilation of migrants and ethnic minorities). We characterize the open-loop Nash equilibrium of the resulting differential dynamic game. In this framework, forward-looking leaders fully take into account the dynamic effects of their cultural action on the evolution of cultural traits in the population. Two important features come out from the analysis. First, the cultural socialization game has some negative sum features for the cultural leaders. As a matter of fact, because increased diffusion of a given cultural trait means decreased diffusion of the other trait in the population, leaders may engage into excessive cultural competition whereby they just annihilate each other. Second, we show that the steady-state cultural equilibrium depends on the time preference structure of the two leaders. Indeed, the fraction of people of a given type increases with the degree of patience of the leader of its own type and decreases with the degree of patience of the leader of the other type. This feature emphasizes how the relative degree of institutionalization and temporal stability of cultural infrastructures (or the cultural leaders that represent them) matter for the resilience of specific minority groups.

The rest of the paper unfolds as follows. In the next section, we relate our paper to the relevant literature. In Section 3, we develop our model with a forward-looking leader and where parents socialize their offsprings. Section 4 is devoted to analysis of the steady-state equilibrium and displays our main results. In Section 5, we look in more detail at the dynamics of our model by studying how a shock affects the short-run and long-run reactions of the leaders in equilibrium. In Section 6, we consider different extensions of the benchmark model. Section 7 studies the case of two leaders who compete against each other. Finally, Section 8 concludes. All proofs are relegated to the Appendix. In the Online Appendix (supplementary material), we propose an alternative formulation of the payoff structure (Online Appendix A), extends the benchmark model in different directions (Online Appendix B), consider the financing of the public good by community members (Online Appendix C), provide a model where the leader(s) directly affects the cultural transmission mechanism (Online Appendix D).

2. Related literature

Cultural transmission and identity/assimilation. Our paper is related to the literature on cultural transmission and, in particular, on the transmission of ethnic identity. Based on some works on anthropology and sociology (see, in particular, Alba, 1990; Bernal and Knight, 1993; Boyd and Richerson, 1985; Cavalli-Sforza and Feldman, 1981; Phinney, 1990), there is a literature in economics initiated by Bisin and Verdier (2000, 2001)[1] arguing that the transmission of a particular trait (religion, ethnicity, identity, social status, etc.) is the outcome of a socialization inside (parents) and outside the family (peers and role models). Our framework also connects to the recent approach on *identity economics*, started by Akerlof and Kranton (2010), which incorporates into economic theory the idea that people do not only pursue economic objectives, but also exert effort to gain and/or retain acceptance into a social group with which they identify. Akerlof and Kranton (2010) provide an extensive list of examples of such phenomena, ranging from schools to street gangs to factories, and showing that when people come from a group that is in some way at odds with the norms of the institutions they participate in, they may act in ways that assert their identity as something different.[2] The phenomenon, originally highlighted by Fordham and Ogbu (1986), of "acting White" is a good example of such an identity choice. Indeed, a large literature has now well documented the fact within inner-city areas, some Black Americans feel that investing in education or getting a mainstream or corporate job is "selling out", a betrayal of one's heritage or culture.[3] This literature does not, however, look at the role of cultural leaders in the formation of identity and in the cultural-transmission process. It mainly focuses on the role of parents, peers and social norms in the identity formation.[4] Our main contribution to this literature is to add this *centralized* and *coordinated dimension* of transmission of

[1] For a review, see Bisin and Verdier (2011).

[2] See Bisin et al. (2011, 2016) and Panebianco (2014) for models of cultural transmission with identity choices.

[3] See, for instance, Ainsworth-Darnell and Downey (1998), Austen-Smith and Fryer (2005), Bisin et al. (2008), Constant and Zimmermann (2008), Fryer and Torelli (2010), Battu and Zenou (2010), Patacchini and Zenou (2016).

[4] An exception can be found in the recent literature on nation-state building, which considers the political incentives for a government to promote a shared national identity in its territory (Alesina and Reich, 2015). Recently, Almagro and Andres (2017) consider a dynamic model of cultural transmission model with a forward-looking government. Their

cultural traits and to analyze how cultural leaders affect the parent's effort in transmitting their trait as well as the identity's choice of ethnic minorities.

Club goods and religion. There is an interesting literature that views religion as a "club good" and look at leaders who can put restrictions on activities, which essentially acts as a tax on outside earnings (Iannaccone, 1992; Berman, 2000; Carvalho and Koyama, 2016).

In our paper, the religious good is also a club good since it is excludable but non-rivalrous. However, it is different to the way it has been modeled in the literature on religious (club) goods. In the latter, there is a fee to access the religious (club) good and the fee is a way to avoid free-riding problems. In our benchmark model, there is no fee. The entry barriers to the religious good are only based on preferences. The action choices between which good to consume act as a cost. Basically, if someone is a secular Muslim, he or she simply does not enjoy going to the mosque and thus obtains zero utility from it. As a result, there is no fee to go to the mosque for secular individuals, they just do not derive any utility from it and thus will not consume this good. The same is true for religious individuals for the consumption of the secular good.

As a result, our model is related to the religious (club) good literature but provides a different angle to this issue (focusing on preferences rather than fees). For example, when Berman (2000) analyzes the issue of orthodox Jews, then his main arguments is that the entry costs are very high so only very motivated Jews will become orthodox Jews. While not focusing on orthodox Jews, in this paper, we argue that preferences can drive individuals to sort into different religious goods. Apart from the cost, individuals who become orthodox Jews must have strong preferences to adhere to this type of behavior. In this respect, our paper connects the two literatures on club goods and cultural transmission.[5]

The role of leaders in influencing culture. As stated in the Introduction, it is well-documented that cultural leaders have an important impact on the identity and cultural choices and integration of immigrants and ethnic minorities in Europe and the United States. There are, however, few economic models that study such issues. Exceptions include Verdier and Zenou (2015), Hauk and Mueller (2015), Carvalho and Koyama (2016) and Prummer and Siedlarek (2017).

As in our paper, Verdier and Zenou (2015) study the transmission of two cultural traits but their main focus is on the interaction between two leaders with opposite objectives. They show that the presence of leaders can prevent the full integration of ethnic minorities. In this analysis however, cultural leaders are only one step forward-looking, the socialization effort of parents is exogenous and each leader chooses the probability of socialization of the children. Our current framework is much more general since we develop a model where leaders are perfectly forward looking, the socialization effort of parents is endogenous and the objective function and choice of the leaders are different since the leaders choose the amount of public good to provide to people from their community while, in Verdier and Zenou (2015), leaders choose the probability of directly affecting children of their own group.[6]

model is able to explain relevant historical case studies where nation-states, facing similar but not completely identical historical contexts, followed disparate nation-building policies.

5 The access cost and preference views on club goods are also connected when the provision of the club good involves a moral hazard problem at the group level and that preferences are not fully observable. In that case, a fee is not only an access cost but also a mechanism of preference signaling (Iannaccone, 1992).

6 In the Online Appendix D, we also consider a model where the objective function of each leader is the same as in Verdier and Zenou (2015) but where leaders are perfectly forward looking and parents choose their socialization effort.

Hauk and Mueller (2015) consider a model of cultural conflict where cultural leaders supply and interpret culture. The authors want to explain the "clash of civilizations" or "clash of cultures" between different religions and to highlight the role of cultural leaders who can amplify disagreement about cultural values. In our model, the leader chooses the level of religious good to provide to the members of his community and has to decide at which moment in time he will do so. So the dynamic aspect of our problem is key to understand when a leader will be active or not. Hauk and Mueller (2015) only focus on steady states and not on the dynamic transition path to such steady states.

Carvalho and Koyama (2016) analyze religious goods by focusing on the trade off between time and money contribution to a religious good. Agents choose whether to join the religious community and divide effort between income-generating activity outside the community and production of a religious club good within the community. The religious authority (the cultural leader) imposes a linear tax on income-generating activity outside the community. This 'tax' to the members of the community corresponds to the severity of prohibitions imposed by the group and is referred to as its level of strictness. They show that, if economic development is sufficiently low, the religious authority (the leader) chooses a strategy of cultural resistance in every period. If economic development is sufficiently high, then the religious authority adopts a strategy of cultural integration in every period. As a result, the focus of Carvalho and Koyama (2016) is very different to ours since they want to explain the impact of Jewish Emancipation and economic development on Jewish religious culture in 19th century Europe. In our model, the tradeoff for the cultural leader is between the parents' socialization efforts, the child's peers and the cost of producing the religious public good. We show that there exists a threshold size in terms of population above which the cultural leader becomes active in terms of producing the religious good.

Finally, Prummer and Siedlarek (2017) propose an interesting model explaining the persistent differences in the cultural traits of immigrant groups with the presence of community leaders. In their model, the leaders influence the cultural traits of their community, which have an impact on the group's earnings. They determine whether a community will be more assimilated and wealthier or less assimilated and poorer. They find that cultural transmission dynamics with two opinion influencers (the host society and the group leader in their setting) result in intermediate long-run integration outcomes for the population under study.

Their model is, however, quite different to ours along several dimensions. First, the parents' role in their child's socialization is not modeled while the child's decision is in terms of identity and skill acquisition. Second, they highlight the network connecting the community while we don't. Third, in Prummer and Siedlarek (2017), there is one cultural leader who chooses the level of identity of his followers. In our model, the leader chooses the level of a cultural good that interacts with family and peer socialization efforts. Also, in Section 7, we model the competition between two leaders with opposite objectives. Finally, Prummer and Siedlarek (2017) only focus on long-run outcomes (steady states) while we also investigate the complete dynamics of the evolution of the fraction of individuals with a certain trait and, in particular, how we discuss how a shock in some exogenous parameter affects this transitional dynamics.

To summarize, Prummer and Siedlarek (2017) consider the role of a leader in an environment where the interaction between workers' productivity and social networks effects matter, while we analyze the role of the leader in an environment where identity choices complement to the provision of a community public good, parents' socialization efforts and peers effects matter. We view our paper and that of Prummer and Siedlarek (2017) as complementary in providing an explanation of the assimilation and identity choices of immigrants.

3. The model

We consider a model of cultural evolution in a *two-cultural trait population* of individuals in which vertical socialization inside the family interacts with oblique socialization outside the family. To be more specific, suppose there are two possible types of cultural traits in the population denoted by a and b (they could refer to "religious" and "secular" trait or "ethnic" and "majority" trait, etc.). The fraction of individuals with trait $i \in \{a, b\}$ is denoted by q^i. Families are composed of one parent and a child, and hence reproduction is a-sexual. All children are born without defined preferences or cultural traits and are first exposed to their parent's trait. *Direct vertical socialization* to the parent's trait, say i, occurs with probability d^i. If a child from a family with trait i is not directly socialized, which occurs with probability $1 - d^i$, she is then subject to *outside socialization*. In that case, the child is matched to a *passive* role model randomly chosen in the society so that she adopts trait i with probability q^i and trait $j \neq i$ with probability $q^j = 1 - q^i$. For $i \in \{a, b\}$ and $j \in \{a, b\}$, denote by P^{ij} the probability that a child from a family with trait i is socialized to trait j. By the Law of Large Numbers, P^{ij} also denotes the fraction of children with a type i parent who have preferences of type j. The socialization mechanism described above can then be characterized by the following transition probabilities, for all $i, j \in \{a, b\}$:

$$P^{aa} = d^a + (1 - d^a)q^a \tag{1}$$

$$P^{ab} = (1 - d^a)(1 - q^a) \tag{2}$$

$$P^{bb} = d^b + (1 - d^b)\left(1 - q^a\right) \tag{3}$$

$$P^{ba} = (1 - d^b)q^a \tag{4}$$

Each individual of type $i \in \{a, b\}$ has preferences $U^i(x, G^i)$, where $x \in \{A, B\}$ is an individual action and G^i is a public good specific to trait i. If, for example, we consider Muslim individuals and interpret the traits a and b as "religious" and "secular", then action $x = A$ means that the individual takes a religious action, i.e. goes often to the mosque, while action B means that she takes a secular action, i.e. never goes to the mosque. Also, in this context, G^a will be a religious public good (a mosque) while G^b could be some secular good such as an historical or cultural centrum on immigration in the host country. We assume the following linear payoff structure:

$$\begin{aligned}
U^a(A, G^a) &= u^a + v^a G^a \\
U^a(B, G^a) &= 0 \\
U^b(B, G^b) &= u^b + v^b G^b \\
U^b(A, G^b) &= 0
\end{aligned} \tag{5}$$

where $u^i, v^i > 0$. This means that, for each individual $i \in \{a, b\}$, her utility is a separable function of $u^i > 0$ and the public good G^i. For each individual of type $i \in \{a, b\}$, the incentives to choose her preferred action x^i increase linearly with the amount of the specific public good provided for that trait. Note that this specification captures, albeit in a simplified way, the idea that there is a complementarity between the preferred action x^i and the specific public good G^i of an individual of type i. Indeed $U^a(A, G^a) - U^a(B, G^a)$ is increasing in G^a while $U^b(B, G^b) - U^b(A, G^b)$ is increasing in G^b. Note also that we assume that $U^a(B, G^a) = U^b(A, G^b) = 0$. This is just a normalization assumption stating that taking an action that is opposite to your own trait gives you zero utility. For example, for a secular person (type b), going to the mosque (action $x = A$) gives no utility.

In the Online Appendix A, we propose a (non-normalized) alternative but equivalent formulation of the payoff structure, which is based on the productivity and discrimination of ethnic minorities. All our results are same under the two formulations. In the main text, we will use payoffs (5) but we will interpret some policy results using payoffs (A.1) from the Online Appendix A since they can give some insights on the productivity and discrimination of ethnic minorities.

We assume that the technology of producing the public good G^i is with constant marginal cost c^i up to a capacity constraint $\overline{G}^i$ so that $G^i \in \left[0, \overline{G}^i\right]$. While capturing the essential features of a convex production technology, this assumption will allow us to fully characterize the transitional dynamics of the socialization mechanism under cultural leadership.

3.1. Population dynamics

The fraction of individuals with trait $i \in \{a, b\}$ at time t is denoted by q_t^i. The process of preference evolution follows an intergenerational model of cultural transmission a la Bisin and Verdier (2001). Noting that $q_t^a = 1 - q_t^b = q_t$ and using (1) and (4), the cultural dynamics of q_t is then given by:

$$\begin{aligned} q_{t+1} &= q_t P^{aa} + (1 - q_t) P^{ba} \\ &= q_t \left[d^a + (1 - d^a) q_t\right] + (1 - q_t)(1 - d^b) q_t \end{aligned}$$

or, equivalently,

$$q_{t+1} - q_t = q_t (1 - q_t) \left(d^a - d^b\right)$$

The continuous-time dynamics version of this equation is given by:

$$\frac{dq_t}{dt} \equiv \overset{\bullet}{q}_t = q_t (1 - q_t) \left(d^a - d^b\right) \tag{6}$$

We assume that the parents of type $i \in \{a, b\}$ are paternalistic and choose their socialization rates d^i in order to maximize the following utility function:

$$V^i = \beta \left[P^{ii} U^i (x, G^i) + P^{ij} U^i (y, G^i)\right] - C(d^i)$$

where $x \in \{A, B\}$ (respectively $y \in \{A, B\}$, with $y \neq x$) is the optimal action of cultural trait $i \in \{a, b\}$ (respectively of the other cultural trait $j \neq i$), $C(d^i) = (d^i)^2/2$ is the quadratic cost of socialization and β is the weight of imperfect altruism that motivates socialization by the parent to transmit her own trait to her offspring. Under these assumptions and using equations (1)–(4) and payoffs (5), the utility function of a parent of type $i \in \{a, b\}$ can be written as:

$$V^a = \beta \left[d^a + (1 - d^a) q^a\right] \left(u^a + v^a G^a\right) - \frac{(d^a)^2}{2}$$

$$V^b = \beta \left[d^b + (1 - d^b) \left(1 - q^a\right)\right] \left(u^b + v^b G^b\right) - \frac{(d^b)^2}{2}$$

At time t, each parent of type $i \in \{a, b\}$ chooses d_t^i that maximizes V^i. We obtain[7]:

$$d_t^i = \beta \left(1 - q_t^i\right)\left(u^i + v^i G_t^i\right) \tag{7}$$

which implies that (noticing that $q_t^a = q_t$):

$$d_t^a - d_t^b = \beta \left[(1 - q_t)\left(u^a + v^a G_t^a\right) - q_t \left(u^b + v^b G_t^b\right)\right]$$

Using this last equation, we can then rewrite the cultural dynamics equation (6) as:

$$\overset{\bullet}{q}_t = \beta q_t (1 - q_t)\left[(1 - q_t)\left(u^a + v^a G_t^a\right) - q_t \left(u^b + v^b G_t^b\right)\right] \tag{8}$$

3.2. *Dynamics of the cultural leader's choice*

As stated above, we assume that group a (say the religious group) has a cultural leader (or institution) that can provide the group specific public good G^a with constant marginal cost $c^a = c$ up to a capacity constraint $\overline{G}^a$ so that $G^a \in \left[0, \overline{G}^a\right]$. To keep the model tractable, we assume that the other group b has no collective authority to produce the public good and therefore we assume for simplicity that $G^b = 0$. This implies that the dynamics of q_t can be written as:

$$\overset{\bullet}{q}_t = \beta q_t (1 - q_t)\left[(1 - q_t)\left(u^a + v^a G_t^a\right) - q_t u^b\right] \tag{9}$$

Let us now study the role of the perfect-forward looking cultural leader of group a. We will start from the long-run cultural situation when the leader is never active ($G^a = 0$), i.e.

$$q(0) = \frac{u^a}{u^a + u^b} \tag{10}$$

We would now like to characterize the condition for the cultural leader to actively intervene in the socialization process of her own group. Notice that the maximum size of group a that can be reached in the case of maximum public good provision $G^a = \overline{G}^a$ is given by:

$$q(\overline{G}^a) = \frac{u^a + v^a \overline{G}^a}{u^a + v^a \overline{G}^a + u^b} \tag{11}$$

We assume that the utility of the cultural leader of group a is given by

$$\int_0^\infty e^{-\rho t}\left(W^a q_t - c\, G_t^a\right) dt \tag{12}$$

where $\rho > 0$ is the discount rate. In (12), the cost of providing the public good G^a is linear and given by $c\, G^a$. In this formulation, quite naturally, the leader's cultural rents $W^a q_t$ increase with the size of the group q_t (i.e. the fraction of the population with trait a) where W^a is the marginal benefit for the leader of type a to have a fraction q_t of the population of type a. The

---

[7] To be consistent with the fact that $d^i \in [0, 1]$ for all values of $q^i \in [0, 1]$ and for $i \in \{a, b\}$, we impose the following restrictions on β, u^i, v^i and $\overline{G}^i$:

$$\beta \max_{i \in \{a,b\}} \left[u^i + v^i \overline{G}^i\right] < 1.$$

cultural rents W^a are the (marginal) benefits for the leader of having a higher fraction of individuals of type a. Think of religious leaders, such as Imams. Obviously, Imams have higher utility when there are more Muslims who become religious. This means that W^a are "ideological" rents. More generally, W^a captures any rents that cultural leaders obtain when they are more individuals in the population who are like them, i.e. who adopt the same values as the leader himself.[8]

The program of the leader can thus be written as:

$$\max_{0 \leq G_t^a \leq \overline{G}^a} \int_0^\infty e^{-\rho t} \left(W^a q_t - c\, G_t^a \right) dt$$

$$s.t. \quad \overset{\bullet}{q}_t = \beta q_t (1 - q_t) \left[(1 - q_t) \left(u^a + v^a G_t^a \right) - q_t u^b \right] \tag{13}$$

$$s.t. \quad q_0 = q(0) \text{ is given}$$

We obtain the present-value Hamiltonian:

$$H_t(G_t^a, q_t, \mu) = e^{-\rho t} \left(W^a q_t - c\, G_t^a \right) + \mu \left\{ \beta q_t (1 - q_t) \left[(1 - q_t) \left(u^a + v^a G_t^a \right) - q_t u^b \right] \right\}$$

where μ is the Lagrangian multiplier (or co-state variable) corresponding to $\overset{\bullet}{q}_t$. It should be noted that this optimization problem is of a bang-bang nature. It corresponds to an epidemiology optimal control problem whereby the leader maximizes a discounted intertemporal payoff function under a "replicator dynamic" equation of the form:

$$\overset{\bullet}{q}_t = \beta q_t (1 - q_t) \left(F^a - F^b \right)$$

where $\overset{\bullet}{q}_t$, the rate of change of the cultural frequency q_t, depends on the strength of the "cultural" selection as reflected by the variance of the dyadic population (the term $q_t(1 - q_t)$) while the bracket term $\left(F^a - F^b \right)$ corresponds to the difference between the "cultural fitness" of trait a (i.e. $F^a = (1 - q_t) \left(u^a + v^a G_t^a \right)$) and the "cultural fitness" of trait b (i.e. $F^b = q_t u^b$). Moreover, given the fact that generally the Hamiltonian $H_t(G_t^a, q_t, \mu)$ is not concave in G_t^a and q_t,[9] the first-order condition approach is not sufficient to characterize optimal trajectories. Fortunately, there is a way out by using another characterization method based on the so-called "Most Rapid Approach Path" formulation of the problem (Spence and Starrett, 1975), which we develop now.

Normalize, for simplicity, $\beta = v^a = 1$. We then have:

[8] Below, in one extension, the leader maximizes parents' socialization rents (see Section 6.2 and the Online Appendix B.2), which means that the cultural rents are related to the fact that parents go to the mosque. In that case, the cultural rents obtained by the leader could be of monetary nature since the more the parents go to the mosque, the more likely they give money to the mosque, which increases the utility of the cultural leader (the Imam in this case).

[9] Indeed, it is easily verified that the Hessian matrix of the Hamiltonian

$$\begin{pmatrix} H_{GG} & H_{Gq} \\ H_{qG} & H_{qq} \end{pmatrix}$$

has a negative determinant and thus cannot be semi-definite negative. The Hamiltonian is therefore not globally concave in (G, q) and the sufficient Magasarian conditions cannot be applied to ensure an optimum solution.

Lemma 1. *Up to some constant, the optimal control problem (13) is equivalent to:*

$$\max_{0 \le G_t^a \le \overline{G}^a} \int_0^\infty e^{-\rho t} R(q_t) dt \tag{14}$$

where

$$R(q_t) = W^a q_t - c\left(\frac{u^b q_t + \rho}{1 - q_t}\right) - c\rho \log\left(\frac{q_t}{1 - q_t}\right)$$

The solution of (14) is a standard optimal control problem in which the control G_t^a is bounded inside an interval $\left[0, \overline{G}^a\right]$ and does not appear anymore in the objective to be maximized. As suggested by Spence and Starrett (1975), the optimal trajectory of this problem takes a "Most Rapid Approach Path" (MRAP hereafter) form, which has the property that it approaches as rapidly as possible some point q^* that locally maximizes the function $R(q_t)$ and stays there forever, given the constraint that q^* can be reached using the control $G^a \in \left[0, \overline{G}^a\right]$. The characterization of point q^* (and the associated control function G_t^{a*}) clearly depends on the shape of the function $R(q_t)$ (its maxima and minima structure) and the degree of feasibility of these points through the control $G^a \in \left[0, \overline{G}^a\right]$.

We obtain then the following lemmas on the shape of the function $R(q_t)$ that will be useful in the sequel.

Lemma 2. *There exists a threshold $\overline{W}$ such that, for $W^a < \overline{W}$, the function $R(q)$ is decreasing in q. For $W^a > \overline{W}$, there exists two interior points q_L and $q_H \in (0, 1)$ such that $R(q)$ reaches a local minimum at q_L and a local maximum at q_H. The threshold $\overline{W}$ is an increasing function of c, ρ and u^b. Also q_L is a decreasing function of W^a, and an increasing function of c, ρ and u^b, while q_H is an increasing function of W^a, and a decreasing function of c, ρ and u^b.*

The shape of the function $R(q)$ is displayed in Fig. 1. As it can be seen for $W^a > \overline{W}$, it has a local minimum at q_L and a local maximum at q_H. There also exists a point $q_S^* \in (0, q_L)$ *such that* $R(q_S^*) = R(q_H)$. Specifically, we have the following lemma.

Lemma 3. *For $W^a > \overline{W}$, there exists a unique point $q_S^* \in (0, q_L)$ such that $R(q_S^*) = R(q_H)$. Moreover, q_S^* is a decreasing function of W^a, and an increasing function of c, ρ and u^b.*

4. Steady-state equilibrium

Let us now analyze the steady-state equilibrium. As stated above, the optimal trajectories of problem (14) take the form of Most Rapid Approach Path (MRAP). Different configurations can be studied. Remember that the initial state is given by the steady-state cultural outcome with no cultural leader at all time (i.e. $G^a(t) = 0$), which is given by (10). Recall, as well, that the long-run steady state when the cultural leader is at maximum capacity at all time (i.e. $G^a(t) = \overline{G}^a$) is given by (11) when $v^a = 1$. The following proposition provides a full characterization of the optimal cultural control trajectory of the leader depending on the initial condition of the population.

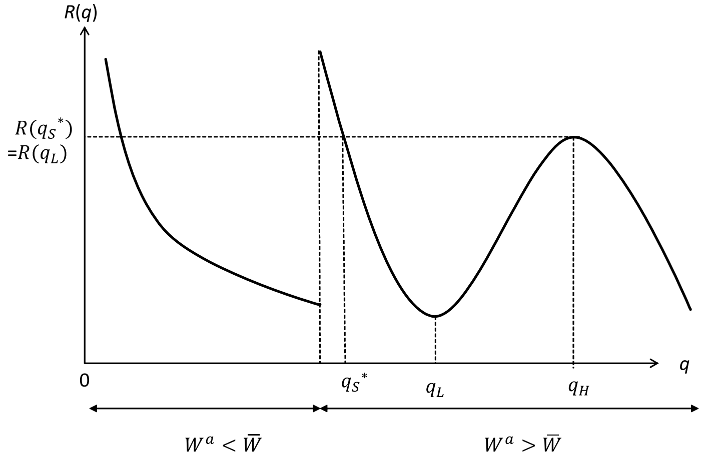

Fig. 1. The function $R(q)$.

Proposition 1.

(*i*) *When $W^a < \overline{W}$, the optimal trajectory for the leader is to remain inactive at all time (i.e. $G^a(t) = 0$) and the cultural steady state stays at $q(0) = u^a / \left(u^a + u^b\right)$ indefinitely.*

(*ii*) *When $W^a > \overline{W}$, there exists a threshold level $\widehat{q}^* \in \left(q_S^*, q_L\right)$ such that:*

- (*ii*1) *If $q(0) < \widehat{q}^*$, the optimal trajectory for the leader is to remain inactive at all time (i.e. $G^a(t) = 0$) and the cultural steady state stays at $q(0) = u^a / \left(u^a + u^b\right)$ indefinitely.*
- (*ii*2) *If $q(0) \in (\widehat{q}^*, q_H)$ the optimal trajectory for the leader is to be active to full capacity (i.e. $G^a(t) = \overline{G}^a$) up to the moment where $q^F = \min\left[q_H, q(\overline{G}^a)\right]$ is reached. Whenever this occurs in finite time T (i.e. when $q_H < q(\overline{G}^a)$), the leader ensures that the cultural steady state stays at $q^F = q_H$ by choosing a "singular" interior policy $G^a(t) = G^F < \overline{G}^a$.*
- (*ii*3) *If $q(0) \geq q_H$, the optimal trajectory is for the leader to remain inactive at all time (i.e. $G^a(t) = 0$) and the cultural steady state stays at $q(0) = u^a / \left(u^a + u^b\right)$ indefinitely.*

Proposition 1 displays two features. First, the cultural leader can only be active when the marginal cultural rent W^a is above a certain threshold $\overline{W}$. Otherwise, she will remain inactive since, by increasing G^a, she can only reduce her net (of cost) cultural rents compared to the level without action. Second, even when $W^a > \overline{W}$, there exists a threshold size, $\widehat{q}^*$, in terms of population with trait a, that needs to be satisfied for the cultural leader to become active. In Fig. 2a, the initial size $q(0)$ of the population of type a is less than the threshold $\widehat{q}^*$. It is then not profitable for the cultural leader to promote more socialization than what vertical socialization by parents of the cultural group already does and the system stays at its initial steady state $q(0)$.

Similarly, when vertical socialization by parents is strong enough to generate a steady cultural state with $q(0) > q_H$ (see Fig. 2d), it does not pay for the cultural leader to promote more cultural transmission to her group. Families already do enough of a good job in transmitting trait a to their offsprings so that the leader does not find it profitable to spend additional resources (the specific public good G^a) to stimulate more cultural transmission.

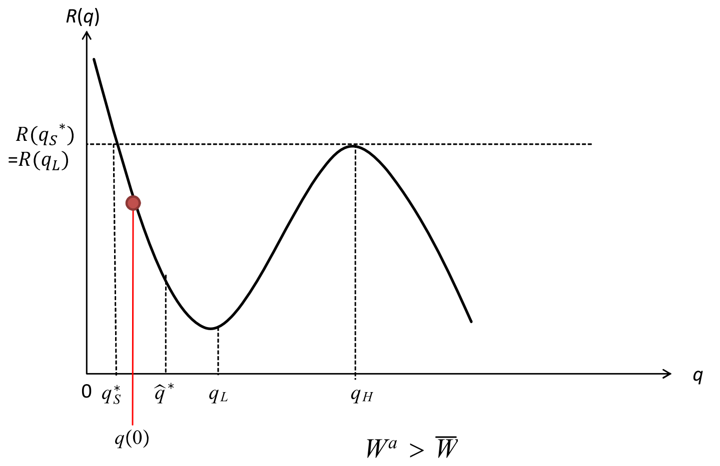

Fig. 2a. Optimal cultural dynamics when $q(0) < \widehat{q}^*$.

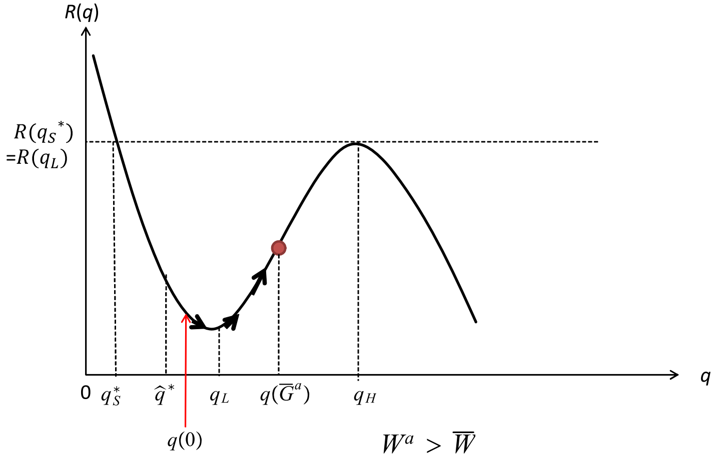

Fig. 2b. Optimal cultural dynamics when $\widehat{q}^* < q(0) < q_L$ and $q(\overline{G}^a) < q_H$.

When the group size is intermediate, as illustrated in Figs. 2b and 2c, i.e. when $q(0) \in (\widehat{q}^*, q_H)$, then the cultural leader is active and pushes forward the cultural dynamics towards the steady state q^F. In such a case, the leader exerts the maximum effort until reaching that steady state. When the steady-state value at full capacity $q(\overline{G}^a)$ is less than q_H (Fig. 2b), then the dynamic system converges asymptotically towards $q^F = q(\overline{G}^a)$ with a cultural leader permanently active at this full capacity. When $q(\overline{G}^a) > q_H$, the cultural steady-state is reached in finite time at $q^F = q_H$ (Fig. 2c), then the leader reduces her effort just to ensure that cultural evolution remains at this optimal long run steady state q_H.

The existence of the critical threshold $\widehat{q}^*$ comes intuitively in the intertemporal tradeoff from the point of view of the cultural leader. Indeed, when the initial size of the population $q(0)$ is less than q_L, the leader's activity G^a stimulates an increase in q. This, in turn, initially involves a short-term rent loss (as long as $q(t)$ stays below q_L, $R(q(t))$ goes down) before eventually generating some higher long-run gains along the upward sloping part of the curve $R(q)$. When

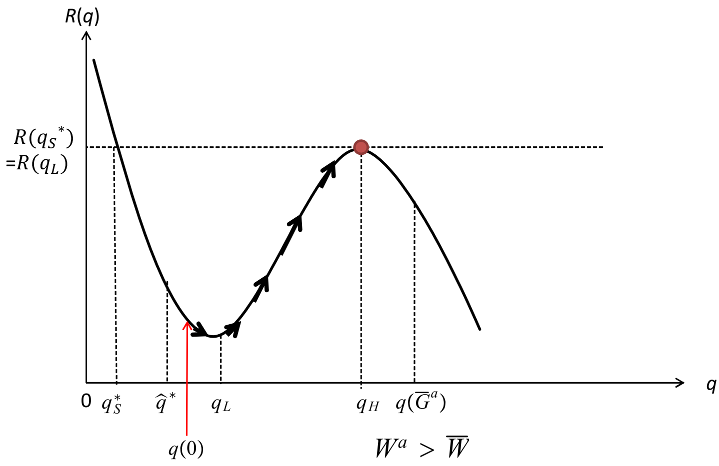

Fig. 2c. Optimal cultural dynamics when $\widehat{q}^* < q(0) < q_L$ and $q(\overline{G}^a) > q_H$.

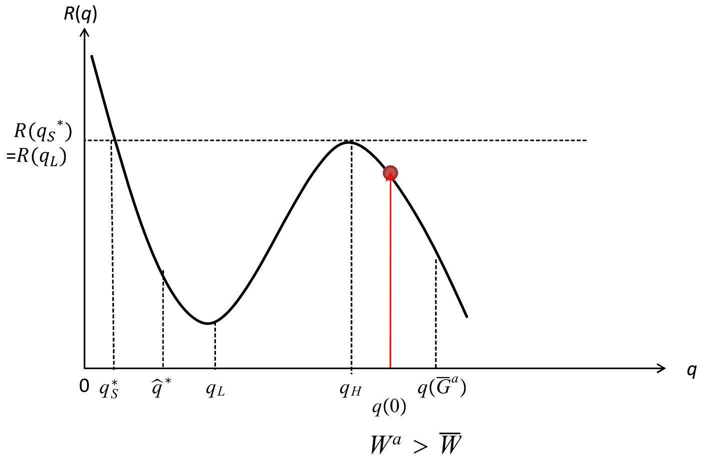

Fig. 2d. Optimal cultural dynamics when $\widehat{q}^* < q(0) < q(\overline{G}^a)$.

deciding whether or not to be active, the forward-looking cultural leader therefore weighs the intertemporal losses of the utility flow $R(q)$ that she first suffers against the long-run gains of eventually a larger steady state population of cultural followers.

When $q(0)$ is close to the minimum value q_L, any public good provision G^a that stimulates positive dynamics of $q(t)$ compares the negligible short-run losses against the significant long-run gains associated to population sizes larger than q_L. Conversely, when $q(0)$ is close to q_S^* (thus further away from q_L), remaining inactive already gives the cultural leader almost the highest possible long-run intertemporal payoff that she can ever obtain with public good provision and cultural activity. At the same time, though, inducing an increase in $q(t)$ with the provision of the public good definitively induces a significant transitory loss of rents for $q(t) \in \left[q_S^*, q_H\right]$. In such a case, the cultural leader has no incentives to start the public good production. It follows by continuity that there is an intermediate size $\widehat{q}^* \in \left[q_S^*, q_L\right]$ such that, for $q(0) < \widehat{q}^*$, there is no provision of public good G^a while, for $q(0) \in (\widehat{q}^*, q_H)$, the cultural leader engages into public good production.

Part (ii) of Proposition 1 highlights the interesting fact that it is only for some intermediate steady-state size $q(0)$ of group a that the cultural leader is active. This feature typically reflects two aspects of the nature of the interactions between purposeful socialization through a leader and the otherwise evolutionary cultural diffusion process occurring through families and peers. First, the more parents and peers from group a are successful in transmitting their trait relative to the other group, the less likely the leader has incentives to actively intervene in the transmission process to sustain the viability of that trait at the group level. This is related to the standard *cultural substituability* effect highlighted by Bisin and Verdier (2001) so that socializing agents (parents and leaders) have an incentive to free ride on the oblique and peer social exposure. This feature explains why for a large enough value of u^a of the parent's socialization incentives and therefore a value of $q(0)$ larger than q_H (part $(ii3)$ of the proposition[10]), the cultural leader does not intervene. Family socialization is already strong enough that it does not pay to the leader to act upon it.

At the same time, though, part $(ii1)$ of Proposition 1 indicates that the leader will also not be active when the parent's socialization incentives u^a are sufficiently low so that the value of $q(0)$ is smaller than the threshold level $\widehat{q}^*$. This suggests, on the contrary, some degree of *complementarity* between family socialization and leader intervention for this range of parameters. Indeed, although the leader's technology of provision of the public good G^a has no fixed cost that could justify the existence of a group threshold effect, there is still a social complementarity between the leader's (centralized) influence and the decentralized population dynamics associated to trait a. In her maximization problem, the only motivation for the leader to provide the public good G^a comes from the fact that she fully internalizes the dynamic process of cultural change of trait a in the population. However, the marginal positive impact of providing the specific public good G^a on the population dynamics of group a is higher, the stronger is the cultural selection effect associated to these dynamics. As it is well known, in evolutionary replicator population dynamics, this selection effect is increasing in the cultural variance of the whole population. In our two-trait population, this is captured by the term $q_t(1-q_t)$, which is close to 0 and increasing in q_t for a small enough minority group of type a. This feature introduces therefore the possibility of a cultural complementarity effect whereby the leader's socialization effort is worth undertaking only when the size q_t of group a is sufficiently large. This is the threshold-effect result obtained in part $(ii2)$ of the proposition.

From this, it turns out that the combination of the aforementioned cultural substituability and complementarity effects determines the full impact of the cultural leader on the dynamics of her group. As a consequence, the leader is going to be active in the socialization process only when the complementarity effect is overcome and the substituability effect is still not too strong, namely when family incentives to socialize their offsprings to trait a are intermediate (i.e. a value of u^a such that $q(0)$ is between $\widehat{q}^*$ and q_H).

Given the importance of the threshold $\widehat{q}^*$ (above which a cultural leader becomes active) for the cultural dynamics, the next proposition offers some comparative statics on $\widehat{q}^*$.

[10] Note that $q(0)$ is increasing in u^a while q_H does not depend on u^a.

Proposition 2. *The type-a population threshold $\widehat{q}^*$ is decreasing in W^a, the marginal benefit of having individuals of type a, and increasing in c, the marginal cost of providing the public good. Furthermore, $\widehat{q}^*$ increases with ρ, the discount rate when ρ is small enough.*

This result shows that the cultural leader is more likely to be active (and thus to provide a large quantity of the public good G^a) when the cultural rents W^a are large enough and the cost of providing G^a is low enough.

The effect of the discount factor ρ is in general ambiguous. Indeed, when ρ becomes larger, the leader discounts more the present than the future and two opposite effects come into play. On the one hand, as he becomes more myopic, the leader does not anticipate the positive impact of G^a on future q_t. Thus, he does not want to intervene because of the immediate cost of providing the public good G^a. This effect tends to increase the participation threshold $\widehat{q}^*$. On the other hand, when ρ increases, the leader also discounts more the value of remaining inactive with a constant cultural population at its initial steady state position. This effect promotes active participation and a reduction of the participation threshold $\widehat{q}^*$. In our set-up, one can show that the first effect dominates the second one when the cultural leader is patient enough, and $\widehat{q}^*$ is increasing with ρ, at low values of ρ.[11]

While the policymaker can separately affect the socialization channels of family and cultural leadership, our analysis suggests that taking into account the interactions between these two channels might actually be quite important to assess the impact of an integration policy. To illustrate this, suppose, for instance, that our minority population is characterized by a family motivation u^a to transmit trait a that is strong enough so that the steady-state fraction $q(0)$ of individuals with that trait is larger than q_H.[12] From Proposition 1, we see that the minority population will not have an active leader promoting the dissemination of trait a. Suppose that to satisfy the cultural integration objective, the policymaker decides to implement a policy directly reducing u^a. By not taking into account the potential existence of a community leader, the policymaker would be lead to expect that such reduction in u^a yields a continuous reduction of the steady state fraction $q(0)$ of individuals with trait a in the population. This is true as long as $q(0)$ is larger than q_H. Below such a level however, one will observe a "cultural resistance" phenomenon: a cultural leader will emerge, promoting the diffusion of trait a in order to maintain the steady-state fraction of individuals with trait a at q_H. As q_H does not depend on u^a, keeping on affecting the family channel of cultural transmission through a reduction of u^a will then be ineffective in "improving" integration. Only targeting directly the leader's incentives (or eventually the marginal effect of the specific public good G^a provided by the leader) can be effective in reducing the fraction of individuals with trait a in the population.

The population size effects highlighted in Proposition 1 also suggest interesting implications about the impact of migration shocks in a minority population and its cultural integration consequences. As in Fig. 2a, consider, for instance, that our minority population has settled at an equilibrium with no cultural leader (with a steady state fraction $q(0)$ of individuals with trait a

[11] Note that the comparative statics for u^b is also in general ambiguous. Indeed with a larger value of u^b, the MRAP goes towards a lower long-run value q^F, which makes it less attractive to the cultural leader. Because of that, the leader is less likely to be active. At the same time, an increase in u^b affects also the speed of cultural change. Indeed, along a MRAP towards the steady-state value q^F, the speed of convergence is reduced. This has an ambiguous effect on the intertemporal value of the leader along the MRAP towards q^F. The overall effect of an increase of u^b on the threshold $\widehat{q}^*$ is, therefore, ambiguous.

[12] Hence $q_H < q(0) < q(\overline{G}^a)$ and $q^F = q_H$.

below $\widehat{q}^*$). Suppose, now, that this community experiences a one time migration shock from the country of origin with mostly individuals with trait a. Just after the migration shock, the new fraction of the population with trait a will increase from $q(0)$ to $q'(0)$ (with $q'(0) > q(0)$). If the migration shock is large enough, the value of $q'(0)$ is above the threshold $\widehat{q}^*$, as in Figs. 2b or 2c, which will lead to the emergence of an active community leader promoting the diffusion of trait a. Again, not taking into account the role of the cultural leader in the community could lead to wrong conclusions about the impact of migration on the process of cultural integration of that community. Indeed, in a world without community leaders or cultural institutions, the *reduction in cultural integration triggered by the migration shock is only temporary*. As a matter of fact, the cultural dynamics after migration is still governed by equation (9) with $G^a = 0$ (no leader) and, as a consequence, the fraction of individuals with trait a converges back to $q\,(0)$.

Now, the possibility of cultural leaders dramatically changes the impact of the migration shock on the steady-state equilibrium. First of all, the migration shock triggers, through the entry of a cultural leader, the institutionalization of the cultural transmission of the differential trait a. Second, the emergence of a centralized mechanism of cultural transmission prevents the integration process to return to the initial state $q\,(0)$ of the community. The entry of the leader will indeed amplify the effect of the migration shock and will lead to a permanently lower integration outcome (i.e. a higher fraction $q_H > q'(0) > q(0)$). Importantly, once such new steady state is reached, any integration policy that targets only the parental transmission incentives through u^a will be ineffective in promoting back cultural integration (as indeed q_H does not depend on u^a).

5. Comparative dynamics and overshooting: short-term versus long-term reaction

Assume now that the leader is active and settles at the high steady state q^F. The model can also provide some interesting comparative dynamics in the steady state and along the transitional path. For this, it is convenient to consider only the case of the singular steady state with $q^F = q_H < q(\overline{G}^a)$. In such a case, the leader in the long run undertakes an interior effort of

$$G^F = -u^a + \frac{q_H u^b}{1 - q_H}$$

with q_H an increasing function of W^a, and a decreasing function of c, ρ and u^b. It follows immediately that:

Proposition 3. *Whenever the leader is active, the steady-state interior effort of the cultural leader $G^F = G^{a*}$ is an increasing function of W^a, a decreasing function of c, ρ, and $u^{a\cdot}$ and an increasing (decreasing) function of u^b when u^b is low enough (large enough).*

Hence, the cultural leader will devote more effort to bias the cultural transmission process when cultural rents W^a are higher, the cost c of the group specific public good is low and the leader is sufficiently forward looking. The fact that the cultural leader effort is decreasing in u^a reflects the idea that there is *cultural substituability* between vertical socialization and the leader's steady state level of intervention, so that the leader of type a intervenes less the more parents of type a exert effort in transmitting their trait relatively to parents of type b. Observe that an increase in the socialization incentives u^b of the other group has, in general, an ambiguous effect on the optimal effort of the leader. Indeed, a higher value of u^b induces a larger socialization rate by individuals of type b.

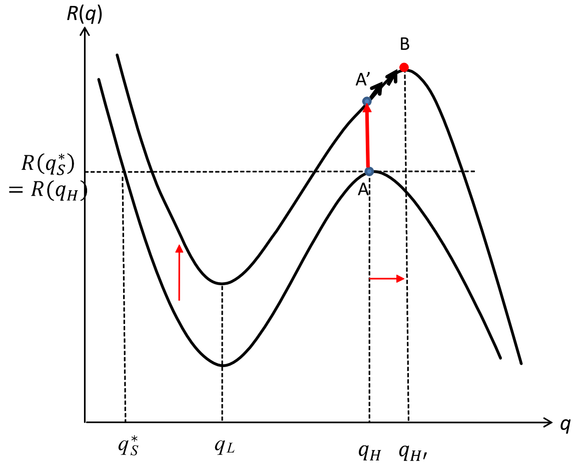

Fig. 3. Overshooting of the leader policy along the transition path after a positive shock on W^a.

Our analysis also allows for some interesting comparative dynamics along the transition path of cultural evolution. Specifically, consider now a positive change in W^a from the steady state (G^{a*}, q_H) depicted by point A in Fig. 3a. As the new steady state $q_{H'}$ is increased from point A to point B following an increase in W^a, the optimal trajectory is a MRAP to the new steady state. As can be seen in the figure, because of an increased value of W^a, the "cultural rent" function of the leader $R(q)$ is first shifted upward to point A'. At this stage, the MRAP involves moving to the long-run steady state (point B). In the short run, this involves a jump in the leader effort from $G^{a*} < \overline{G}^a$ to full capacity $\overline{G}^a$ until the new steady state is reached. If this occurs in finite time (i.e. $q_{H'} < q(\overline{G}^a)$), then the optimal effort of the cultural leader jumps downward to a level $G^{a*\prime}$, which is less than the full capacity $\overline{G}^a$ but larger than the steady-state cultural effort G^{a*} before the shock on W^a. There is therefore a *short-term overshooting* in terms of the cultural leader activity. The resulting long-run effect on the size of the group is positive since $q_{H'} > q_H$.

Consider now a positive shock on the cost parameter c (or the discount factor ρ) of the leader. By a similar argument, the long-run steady state is shifted from point A in Fig. 4 to point B associated to lower long-run size $q_{H'}$ of the cultural group. In the transition, the new trajectory is a MRAP towards this lower value of q_H. This implies that the cultural effort of the leader goes to $G^a = 0$ until the size of the group reaches $q_{H'} < q_H$. When this is achieved, the leader's effort jumps up to a positive effort G^{a*}, which is less than the one exerted at point A before the shock. Hence a positive shock on the cost parameter induces a *short-term undershooting* of the cultural leader's socialization effort.

These transitory dynamics have important policy implications in terms of the reaction of the minority community to changes in the environment of this group. Typically, there will always be some over-reactions or under-reactions compared to the long-run effect that can be expected. As a result, a successful policymaker acting on the parameters of the model (such as W^a or c) should be enough patient to fully see the impact of her policy on the steady-state outcomes.

6. Extensions

In this section, we discuss various extensions of our benchmark model and show that essentially the same qualitative results hold in these contexts.

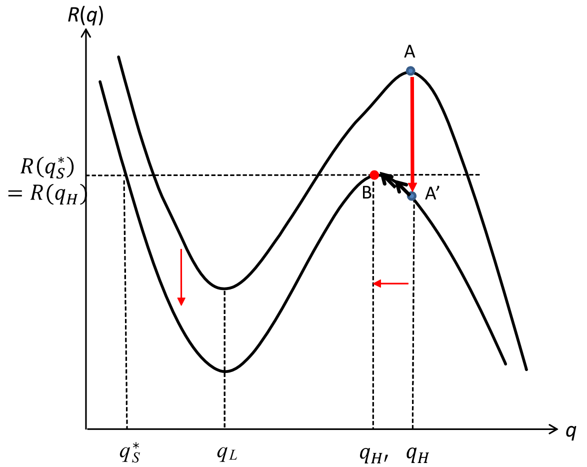

Fig. 4. Undershooting of the leader policy along the transition path after a positive shock on cost c.

6.1. Spillovers of the group-specific public good on individuals from the other group

So far, we have considered a model where the community good G^a only provides benefits to individuals from the same group, i.e. those endowed with preferences of type a. However, the good G^a could also have some *non-rivalry* dimension since individuals from the other cultural trait b could also benefit from G^a. In the Online Appendix B.1, we extend our benchmark model to capture these spillover effects and analyze the implication of the non-exclusivity of the community public good on the steady-state equilibrium.

6.2. Other objective functions of the cultural leader

We assumed that the leader was maximizing a rent proportional to the size of her group minus the cost of the consumption of the public good, i.e. $\int_0^\infty e^{-\rho t}\left(W^a q_t - cG_t^a\right)dt$ (see (13)). We consider two extensions where the leader maximizes a different objective function. First, in the Online Appendix B.2, we look at the case when the leader maximizes socialization rents, which are the fraction of the population of type a times the optimal effort of type-a parents so that her objective function is given by: $\int_0^\infty e^{-\rho t}\left(q_t d_t^a - cG_t^a\right)dt$. Second, in the Online Appendix B.3, we consider the case when the objective function of the leader is to maximize the expected utility of individuals of the same type (i.e. of type a), i.e. the leader maximizes $\int_0^\infty e^{-\rho t}\left[q_t\left(u^a + G_t^a\right) - cG_t^a\right]dt$.

6.3. Financing of the public good by the community

In our benchmark model, we assumed that the cultural leader was financing himself the cost of the provision of the group-specific public good. In particular, in the public good literature on religion (Iannaccone, 1992; Berman, 2000; Carvalho and Koyama, 2016; Carvalho, 2018), the community that accesses the public good also pays for it in some way, either in terms of income or in terms of time. In the Online Appendix C, we extend the basic model to capture these features and allow the public good $G = G^a$ to be financed through a contribution $\tau^a \in [0, \overline{\tau}]$ that community members pay if they consume the good G. For that, in (C.1), we introduce the tax τ^a in the utility function of all individuals who consume good G and determine a budget constraint (C.2) that equalizes the costs and benefits of consuming the public good, i.e. $cG_t = \tau_t^a q_t$. Moreover we allow the leader not only to care about his cultural rents $W^a q_t$ but also to care about the welfare of individuals belonging to the club.

6.4. The cultural leader directly affects cultural socialization

Finally, in the Online Appendix D1, we consider a model where the leader *directly* affects the cultural transmission mechanism. For that, we assume that children are first exposed to their parent i's trait (with probability d^i) but, when this fails, the child is subject to *outside socialization* so that, with probability γ, she is directly exposed to the leader (of trait a) while, with probability $1-\gamma$, the child is matched to a *passive* role model randomly chosen in the society (i.e. she adopts trait i with probability q^i). Again, we show that the main results of the analysis remain the same. In particular, Proposition D1 in the Online Appendix D provides very similar results to the case where the leader chooses the public good G^a instead of direct socialization γ.

7. Competition between cultural leaders

7.1. The model

Let us extend our benchmark model by considering two cultural leaders (for instance, a religious leader and the host-country institution) who compete with each other for transmitting their own trait.[13] As above, the utility of a cultural leader of group $i \in \{a, b\}$ is equal to: $V^i = \int_0^\infty e^{-\rho^i t}\left(W^i q_t - c\, G_t^i\right) dt$, where leader a's cultural rents are $W^a q_t$ (and $W^b(1-q_t)$ for the leader of type b) increase with the size of his own group $q_t^a = q_t$ (resp. $q_t^b = 1-q_t$) and cG_t^a (resp. cG_t^b) is the linear resource cost for the public good.

At time t, each parent of type $i \in \{a, b\}$ chooses d_t^i that maximizes V^i. As in the benchmark model, we obtain (see (7)):

$$d_t^i = \beta\,(1-q_t)\left(u^i + v^i G_t^i\right) \tag{15}$$

which implies that:

$$d_t^a - d_t^b = \beta\left[(1-q_t)\left(u^a + v^a G_t^a\right) - q_t\left(u^b + v^b G_t^b\right)\right] \tag{16}$$

Using this last equation, the cultural dynamics equation (6) can be written as (8), that is:

$$\overset{\bullet}{q}_t = \beta q_t\,(1-q_t)\left[(1-q_t)\left(u^a + v^a G_t^a\right) - q_t\left(u^b + v^b G_t^b\right)\right] \tag{17}$$

This setting describes a *dynamic differential game* of public good choice between the two leaders. As it is well-known, various equilibrium concepts can be used to analyze such games. In the following, we characterize the *open-loop Nash equilibrium* concept. Denote by $G^a(.)$ and $G^b(.) \in \left[0, \overline{G}\right]$ some admissible public good choices of the leaders a and b, and by $V^a\left(G^a(.), G^b(.)\right)$ and $V^b\left(G^a(.), G^b(.)\right)$ the intertemporal values associated to these public good choices, i.e.,

[13] In the Online Appendix D.2, we also develop a model of competition between two forward-looking leaders but using a framework where the leader *directly* affects the cultural transmission (as in Section 6.4 and in the Online Appendix D.1).

$$V^a\left(G^a(.),G^b(.)\right)=\int_0^\infty e^{-\rho^a t}\left[W^a q\left(t,G^a,G^b\right)-c\,G^a(t)\right]dt$$

$$V^b\left(G^a(.),G^b(.)\right)=\int_0^\infty e^{-\rho^b t}\left[W^b\left(1-q\left(t,G^a,G^b\right)\right)-c\,G^b(t)\right]dt$$

where $q\left(t,G^a,G^b\right)$ is the trajectory starting at q_0 and satisfying the differential equation (17).

Definition 1. *An open-loop Nash equilibrium $\left(G^{a*},G^{b*}\right)$ is characterized by the following conditions:*

$$V^a\left(G^{a*}(.),G^{b*}(.)\right)\geq\left(G^a(.),G^{b*}(.)\right) \textit{ for all admissible } G^a(.)$$
$$V^b\left(G^{a*}(.),G^{b*}(.)\right)\geq V^b\left(G^{a*}(.),G^b(.)\right) \textit{ for all admissible } G^b(.)$$

For an open-loop Nash equilibrium in which the cultural leader b chooses $\left(G_t^b\right)_{t\in[0,\infty]}$, the problem of the cultural leader a can be written as:

$$\max_{0\leq G_t^a\leq\overline{G}^a}\int_0^\infty e^{-\rho^a t}\left(W^a q_t-c\,G_t^a\right)dt$$
$$s.t.\quad \overset{\bullet}{q}_t=\beta q_t\left(1-q_t\right)\left[\left(1-q_t\right)\left(u^a+v^aG_t^a\right)-q_t\left(u^b+v^bG_t^b\right)\right] \tag{18}$$
$$s.t.\quad q_0=q(0)\text{ is given}$$

As in the benchmark model, the payoff of leader a is linear G_t^a and because of the shape of the cultural evolution equation (17), the Hamiltonian is not concave in G^a. The first order approach may therefore again be problematic. Moreover, here, we have a differential game between two leaders and we need to keep track of two trajectories at the same time and, therefore, we cannot totally characterize the steady-state solutions when G_t^a and G_t^b vary over time. Still under some specific circumstances, we can still use the Most Rapid Path Approach (MRAP) to highlight how the cultural dynamics and the steady-state equilibria can be affected by cultural leaders' competition.[14] As matter of fact, we will solve the open-loop Nash equilibrium when each cultural leader $i\in\{a,b\}$ only makes constant choices, that is either $G^i=0$ or $G^i=\overline{G}$. In that case, the typical $R(q_t)$ function that we used in the benchmark model to determine the behavior of a given cultural leader i, will only take two possible shapes when the opposite leader $j\neq i$ makes a constant choice $G^j=0$ or $G^j=\overline{G}$. In such a case, we will be able to characterize the open-loop best-reply function of leader i as a MRAP with as well a constant choice for G^i ($=0$ or $\overline{G}$). This will allow us to characterize the open loop Nash equilibria and cultural trajectories under these circumstances.

To be more precise, let us solve the problem of leader a given by (18) and assume that leader b only chooses either $G^b=0$ or $G^b=\overline{G}$. Assume, for simplicity, that $\beta=v^a=v^b=1$ and denote $u^b\left(G^b\right):=u^b+v^bG^b=u^b+G^b$. Then, (18) is equivalent to:

[14] When G_t^a and G_t^b vary over time, the $R(q_t)$ function that we use for the MRAP approach and which is crucial for characterizing the steady-state solutions, is much more complex and in particularly not stationary in time.

$$\max_{0\leq G_t^a \leq \overline{G}^a} = \int_0^\infty e^{-\rho^a t} R(q_t, G^b) c dt + c\left[\log\left(\frac{q_0}{1-q_0}\right) + \frac{1}{1-q_0}\right] + \frac{u^a}{\rho^a} \tag{19}$$

where

$$R(q_t, G^b) = \frac{W^a}{c} q_t - \left[\frac{u^b\left(G^b\right) q_t + \rho^a}{1-q_t}\right] - \rho^a \log\left(\frac{q_t}{1-q_t}\right) \tag{20}$$

As stated above, $R(q_t, G^b)$ takes only two possible shapes: if leader b chooses $G^b = 0$, then $u^b\left(G^b\right) = u^b$ and if leader b chooses $G^b = \overline{G}$, then $u^b\left(G^b\right) = u^b + \overline{G}$.

Consider now the choice of leader b and assume that leader a chooses either $G^a = 0$ or $G^a = \overline{G}$. In that case, the leader b solves the following program:

$$\max_{0\leq G_t^b \leq \overline{G}^b} \int_0^\infty e^{-\rho^b t}\left(W^b(1-q_t) - c\, G_t^b\right) dt$$

$$s.t. \quad \overset{\bullet}{q}_t = \beta q_t\,(1-q_t)\left[(1-q_t)\left(u^a + v^a G_t^a\right) - q_t\left(u^b + v^b G_t^b\right)\right]$$

$$s.t. \quad q_0 = q(0) \text{ is given}$$

Assume again, for simplicity, that $\beta = v^a = v^b = 1$ and denote $u^a\,(G^a) := u^a + v^a G^a = u^a + G^a$ and $\underline{q}_t := 1 - q_t$. Then, this is equivalent to:

$$\max_{0\leq G_t^b \leq \overline{G}^b} \int_0^\infty e^{-\rho^b t}\left(W^b \underline{q}_t - c\, G_t^b\right) dt$$

$$s.t. \quad \overset{\bullet}{\underline{q}}_t = \underline{q}_t(1-\underline{q}_t)\left[\left(1-\underline{q}_t\right)(u^b + G_t^b) - \underline{q}_t\left(u^a + G^a\right)\right] \tag{21}$$

$$s.t. \quad \underline{q}_0 = 1 - q(0) \text{ is given}$$

By comparing (18) and (21), we see that the problem of leader a and that of leader b are totally symmetric. So, when we solve the program for leader a, we can deduce the solution for leader b, by keeping in mind that $\underline{q}_t := 1 - q_t$. For leader b, (21) is equivalent to:

$$\max_{0\leq G_t^b \leq \overline{G}^b} \int_0^\infty e^{-\rho^b t} R(\underline{q}_t, G^a) c dt \tag{22}$$

where

$$R(\underline{q}_t, G^a) = \frac{W^b}{c}\underline{q}_t - \left(\frac{u^a\,(G^a)\,\underline{q}_t + \rho^b}{1-\underline{q}_t}\right) - \rho^b \log\left(\frac{\underline{q}_t}{1-\underline{q}_t}\right)$$

where, again, $R(\underline{q}_t, G^a)$ takes two possible shapes depending on whether $G^a = 0$ or $G^a = \overline{G}$.

In this framework, we will consider two cases: (i) $\rho^a = \rho^b = \rho$, i.e. the two leaders have the same discount rate, (ii) $\rho^a < \rho^b$, i.e. leader a is more patient than leader b.

We focus on the symmetric case when $u^a = u^b = u$ and $W^a = W^b = W$. In this case, the parent' socialization efforts can be written as:

$$d_t^a = (1 - q_t)\left(u + G_t^a\right) \tag{23}$$

$$d_t^b = q_t \left(u + G_t^b\right) \tag{24}$$

which implies that:

$$\begin{aligned} d_t^a - d_t^b &= (1 - q_t)\left(u + G_t^a\right) - q_t\left(u + G_t^b\right) \\ &= \left(u + G_t^a\right) - q_t\left(2u + G_t^a + G_t^b\right) \end{aligned}$$

We will analyze this problem by starting at $q(0) = \frac{u^a}{u^a + u^b} = 1/2$, i.e. the steady state when both leaders are inactive at $G^a = G^b = 0$. Observe that it is also the steady state when the two leaders are both fully active at $G^a = G^b = \overline{G}$.

7.2. The two leaders have the same discount rate

Consider the problem of leader a (the problem of leader b is similar since the two programs are symmetric) and remember that leader b can only make two choices: $G^b = 0$ or $G^b = \overline{G}$. When $u^a = u^b = u$ and $W^a = W^b = W$, (20) can be written as:

$$R\left(q, U, \rho, \frac{W}{c}\right) := R(q_t) = \frac{W}{c}q - \left(\frac{U\,q + \rho}{1 - q}\right) - \rho \log\left(\frac{q}{1 - q}\right) \tag{25}$$

where U is defined as follows: when the leader b chooses permanently $G^b = 0$, then $U = u$, whereas when leader b chooses permanently $G^b = \overline{G}$, then $U = u + \overline{G}$. The following lemma (whose proof is identical to that of Lemma 2) gives the properties of the function $R\left(q, U, \rho, \frac{W}{c}\right)$ defined in (25).

Lemma 4. *There exists a threshold $\overline{W}(\rho, U)$ such that, for $\frac{W}{c} < \overline{W}(\rho, U)$, the function $R\left(q, U, \rho, \frac{W}{c}\right)$ is decreasing in q. For $W > \overline{W}(\rho, U)$, there exists two interior points q_L and $q_H \in (0, 1)$ such that $R\left(q, U, \rho, \frac{W}{c}\right)$ reaches a local minimum at q_L and a local maximum at q_H. The threshold $\overline{W}(\rho, U)$ is a increasing function of ρ and U. Also $q_L = q_L(W/c, \rho, U)$ is a decreasing function of W/c, and an increasing function of ρ and U, while $q_H = q_H(W/c, \rho, U)$ is an increasing function of W/c, and a decreasing function of ρ and u. Finally the function $R\left(q, U, \rho, \frac{W}{c}\right)$ is decreasing in U.*

Consider then the following parameter restrictions:

$$\frac{W}{c} > \overline{W}\left(\rho, u + \overline{G}\right) \tag{26}$$

$$q_L\left(W/c, \rho, u + \overline{G}\right) < \frac{u}{2u + \overline{G}} < \frac{1}{2} < \frac{u + \overline{G}}{2u + \overline{G}} < q_H\left(W/c, \rho, u + \overline{G}\right) \tag{27}$$

These conditions are satisfied when u and $\overline{G}$ are not too large and W/c is large enough. These two conditions together with Lemma 4 ensure that the two functions $R\left(q, U, \rho, \frac{W}{c}\right)$, for $U = u, u + \overline{G}$, have exactly the same shape as the function $R(q)$ depicted in Fig. 1, which is characterized by Lemma 2. We have then the following result:

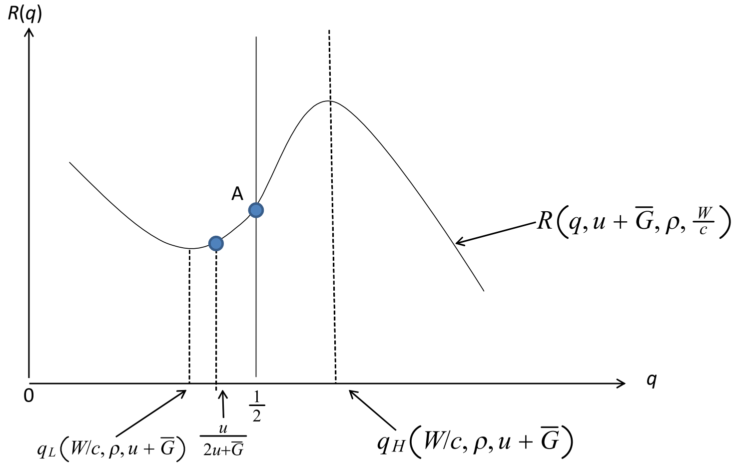

Fig. 5. Open loop Nash equilibrium with two leaders with the same discount rate ρ.

Proposition 4. *Assume that (26) and (27) hold and that both leaders have the same discount rate $\rho^a = \rho^b = \rho$. Then, there exists a unique open-loop Nash equilibrium such that: $G_t^{a*} = G_t^{b*} = \overline{G}$, $\forall t \geq 0$. In that case, the cultural dynamics stay at $q_t = q(0) = 1/2$ at all time and parents' socialization effort is the same and equal to: $d_t^{a*} = d_t^{b*} = \left(u + \overline{G}\right)/2$, $\forall t \geq 0$.*

This situation is illustrated in Fig. 5 where we depict the function $R\left(q, u+\overline{G}, \rho, \frac{W}{c}\right)$ for a given leader $i \in \{a, b\}$ when conditions (26) and (27) are satisfied. When the other leader $j \neq i$ plays permanently $G^j = \overline{G}$, leader i then prefers to stay at point A associated to a steady state $1/2$ with permanent effort $G^i = \overline{G}$, rather than any other trajectory associated with values of $q_t \in \left[\frac{u}{2u+\overline{G}}, \frac{1}{2}\right[$ that would be obtained by playing some value $G_t^i < \overline{G}$. Consequently, the open loop best response function of each leader to the other leader playing a constant effort $G^j = \overline{G}$ is to play $G^i = \overline{G}$ and we obtain the result that there exists an open-loop Nash equilibrium in which both leaders choose constant values of the public good $G_t^{a*} = G_t^{b*} = \overline{G}$, $\forall t \geq 0$. Moreover, under the assumption (27), we show in Lemma 6 in the proof of Proposition 4, we show that playing $G^i = \overline{G}$ is a *dominant strategy* for any leader i when the other leader j is playing permanently a constant value $G^j \in \left[0, \overline{G}\right]$. This implies that $G_t^{a*} = G_t^{b*} = \overline{G}$ is the *unique* open loop Nash equilibrium such that both leaders choose constant values of the public good.

More generally, this proposition illustrates a situation in which both leaders undertake in equilibrium some very tough cultural competition although this does not change at all the evolution of cultural traits in the population (which stays at $q(0) = 1/2$ for ever). The cultural game has some clear negative features. Indeed, when a leader increases the size of its group, it generates a negative externality on the other leader. As the other leader wants to compensate for this by providing more public goods to his own community, this leads to the existence of a Nash equilibrium where both leaders exert a maximal socialization effort equal to $G_t^{a*} = G_t^{b*} = \overline{G}$. Obviously in this context, both leaders would be better off not undertaking any socialization and committing themselves to a policy path without intervention, i.e. $G^a = G^b = 0$. However, such a policy path is not individually incentive compatible. Interestingly, we obtain this result when the family socialization incentives u and the leaders' capacity to affect the cultural dynamics $\overline{G}$ are not too large compared to the reward of socialization W/c, which is guaranteed by (26).

In this case, we may have excess cultural competition between two leaders annihilating each other.

7.3. One leader is more patient than the other

Consider now the case when leader a is more patient than leader b, i.e. $\rho^a < \rho^b$. This case illustrates the role of differential leaders' time discount factors for the dynamics of cultural transmission. Specifically, consider the following parameter restrictions:

$$\overline{W}\left(\rho^a, u\right) < \frac{W}{c} < \overline{W}\left(\rho^b, u + \overline{G}\right) \tag{28}$$

These inequalities guarantee that leader b (who is sufficiently impatient since she has a high enough discount factor ρ^b) will remain inactive, i.e. choose $G_t^b = 0$, $\forall t$, when leader a is choosing the MRAP $G_t^a = \overline{G}$, $\forall t$. We have the following result.

Proposition 5. *Assume that (27) and (28) hold and leader a is more patient than leader b, i.e. $\rho^a < \rho^b$. Then, then there exists a unique open-loop Nash equilibrium such that $G_t^{a*} = \overline{G}$ and $G_t^{b*} = 0$, $\forall t \geq 0$. In such a case, the cultural dynamics converge to*

$$q^* = \frac{u + \overline{G}}{2u + \overline{G}} > 1/2,$$

which is decreasing in u and increasing in $\overline{G}$, while steady state parents' socialization efforts are the same and equal to:

$$d^{a*} = d^{b*} = \frac{u\left(u + \overline{G}\right)}{2u + \overline{G}}$$

which is increasing in both u and $\overline{G}$.

The situation is illustrated in Figs. 6a and 6b where we depict the function $R\left(q, u, \rho^a, \frac{W}{c}\right)$ for leader a and $R\left(\underline{q}, u + \overline{G}, \rho^b, \frac{W}{c}\right)$ for leader b when condition (28) is satisfied and $\underline{q} = 1 - q$. In Fig. 6a, leader a's open-loop best response to player b playing a constant effort $G^b = 0$ is to play the MRAP associated to point A with a steady state $q^* = \frac{u+\overline{G}}{2u+\overline{G}}$ reached through a constant policy $G^a = \overline{G}$. Conversely, in Fig. 6b, when leader a plays permanently $G^a = \overline{G}$, given that, under (28), $R\left(\underline{q}, u + \overline{G}, \rho^b, \frac{W}{c}\right)$ is decreasing in $\underline{q}$, leader b clearly prefers to go to point B associated to the steady state $\underline{q} = 1 - q = \frac{u}{2u+\overline{G}}$ with permanent effort $G^b = 0$, rather than any other trajectory associated with values of $\underline{q}_t \in \left[\frac{u}{2u+\overline{G}}, \frac{1}{2}\right[$. As a consequence, there exists an open-loop Nash equilibrium such that $G_t^{a*} = \overline{G}$ and $G_t^{b*} = 0$, $\forall t \geq 0$. Again, when also (27) holds, $G^a = \overline{G}$ is a *dominant strategy* for leader a to the leader b playing permanently a constant value $G^b \in \left[0, \overline{G}\right]$. In such a case, $G_t^{a*} = \overline{G}$ and $G_t^{b*} = 0$, $\forall t \geq 0$ is the unique the open loop Nash equilibrium such that both leaders choose constant values of the public good.

Proposition 5 shows that the difference of patience between the two cultural leaders may lead to an outcome where only the most patient one is actively promoting the cultural transmission of her trait. This occurs when condition (28) is satisfied. Given that $\rho^a < \rho^b$ and that the

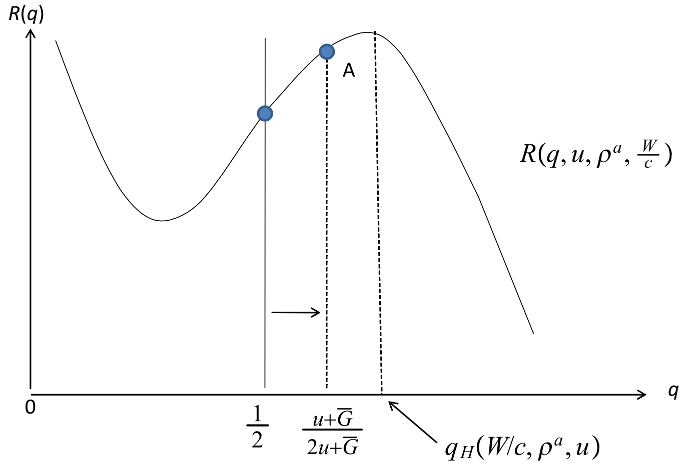

Fig. 6a. Open loop best response for leader a to leader b playing $G^b = 0$.

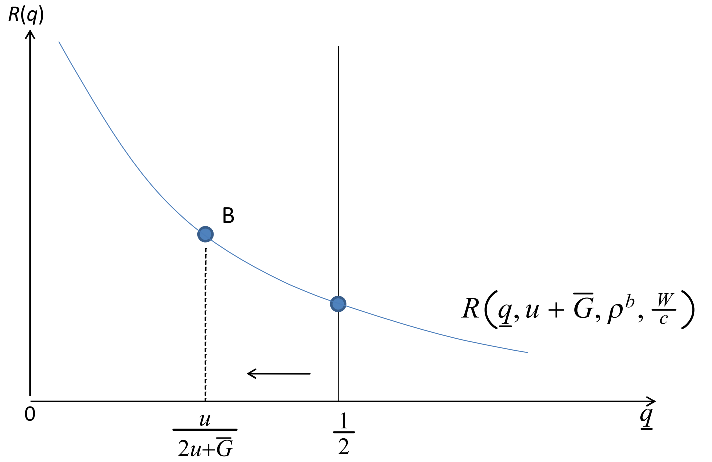

Fig. 6b. Open loop best response for leader b to leader a playing $G^b = \overline{G}$.

threshold function $\overline{W}(\rho, u)$ is increasing in ρ, this condition is compatible with $\overline{W}\left(\rho^b, u\right) < \frac{W}{c}$ which, in fact, would ensure leader b to be active if he were alone to promote the cultural socialization of his own group. The presence of a more patient cultural leader a, stimulating the cultural transmission of the other trait in society, makes it, however, too costly for leader b to intervene. As a consequence, the cultural trajectory converges towards a steady state $q^* > 1/2$, favoring the trait of the most patient leader. Also, an overall reduction in the importance of families as agents of socialization (i.e. a lower value of u) will make it more difficult for the trait associated to the more impatient leader to persist in the long run in the community.

It might be useful to better understand the assumption $\rho^a < \rho^b$ by associating the degree of patience of leaders to the degree of stability of the institution they represent. Suppose, indeed, that leader a (or the institution he represents) promotes "the mainstream" trait of society as an objective of cultural integration of a given minority population and that leader b promotes a differential ethnic trait (say religious or ethnic trait) of that given minority group. Then, it seems natural to assume that leader a is more forward looking than leader b (i.e. $\rho^a < \rho^b$). Indeed, the objective of leader a is similar to the intertemporal objective of the society at large while, that of leader b, reflects the cultural identity of a much smaller group.[15]

In a political economy context, one may also think, however, that the more patient institution (associated to trait a) is the one reflecting strong religious commitments of a minority group, not much subject to strong political competition while the mainstream trait of society (now trait b) is represented by a secular political leader subject to some higher degree of political uncertainty (associated to a probability of local reelection for instance). The previous result suggests that if the political uncertainty of being reelected is large enough, this political leader might refrain from actively implementing a cultural policy against the more "patient" religious leader (i.e. promoting actively the secular trait). Interestingly, in that case, an overall reduction in the importance of families as agents of socialization would makes it easier for the religious trait to persist in the long run.

The general conclusion of this section is that, if two perfectly-forward looking leaders with opposite views and values compete in transmitting their values, then the long-run equilibrium cultural structure of the population will strongly depend on the dynamic incentives of the leaders and on the effort that each type of parent exerts in socializing their offsprings. In particular, the *relative* forward lookingness of cultural leaders or institutions is an important factor affecting the success of the trait they promote. Hence political events (such as competitive elections), affecting asymmetrically the degree of political uncertainty facing different cultural leaders, will, everything been equal, favor the leader securing best her power position in the institution she represents. This highlights an important political economy dimension connected to the issue of cultural integration dynamics. When socialization involves cultural institutions that strategically compete in the "market" for cultural values, then the power structure inside such institutions will matter for the overall cultural dynamics. Cultural institutions with relatively more stable and encompassing and consistent long-term views will have an advantage in promoting their cultural traits.

8. Conclusion

In this paper, we study the population dynamics of cultural traits in a model of intergenerational transmission with perfectly-forward looking cultural leaders and institutions who compete or not for oblique socialization. We emphasize different facets of the impact of cultural leaders in the process of cultural assimilation of minority groups.

In terms of policy implications, our analysis highlights two main aspects. First, the policymaker should take into account the crucial interaction between the *centralized* transmission of a particular trait by leaders and the *decentralized* transmission of the same trait by parents and peers. Second, the policymaker should be patient enough because the *short-term* effects may be very different from the *long-term* ones due to over-reactions or under-reactions of the cultural group in the short-run compared to the long-run.

[15] By definition, although the nature of "mainstream" cultural institutions may change over time, they are more likely to exist in the long run than minority or marginalized cultural group's institutions.

Appendix A. Proofs

Proof of Lemma 1. From (9), we easily obtain:

$$G_t^a = \frac{\overset{\bullet}{q}_t}{\beta v^a q_t (1-q_t)^2} + \frac{q_t u^b}{v^a (1-q_t)} - \frac{u^a}{v^a}$$

Therefore, we have:

$$W^a q_t - c G_t^a = W^a q_t - c\left[\frac{\overset{\bullet}{q}_t}{\beta v^a q_t (1-q_t)^2} + \frac{q_t u^b}{v^a (1-q_t)} - \frac{u^a}{v^a}\right]$$
$$= W^a q_t - \frac{cu^b}{v^a}\frac{q_t}{(1-q_t)} + \frac{cu^a}{v^a} - \frac{c}{\beta v^a}\frac{\overset{\bullet}{q}_t}{q_t(1-q_t)^2}$$

As a result, the program (13) can be written as:

$$\max_{0\le G_t\le \overline{G}} \int_0^\infty e^{-\rho t}\left(W^a q_t - \frac{cu^b}{v^a}\frac{q_t}{(1-q_t)} + \frac{cu^a}{v^a} - \frac{c}{\beta v^a}\frac{\overset{\bullet}{q}_t}{q_t(1-q_t)^2}\right)dt$$
$$= \max_{0\le G_t\le \overline{G}} \left\{\int_0^\infty e^{-\rho t}\left(W^a q_t - \frac{cu^b}{v^a}\frac{q_t}{(1-q_t)} + \frac{cu^a}{v^a}\right)dt - \frac{c}{\beta v^a}\int_0^\infty e^{-\rho t}\frac{\overset{\bullet}{q}_t}{q_t(1-q_t)^2}dt\right\}$$

Observe that:

$$\int_0^\infty e^{-\rho t}\frac{\overset{\bullet}{q}_t}{q_t(1-q_t)^2}dt = \int_0^\infty e^{-\rho t}\overset{\bullet}{q}_t\left[\frac{1}{q_t} + \frac{1}{(1-q_t)^2} + \frac{1}{(1-q_t)}\right]dt$$
$$= -\left[\log\left(\frac{q_0}{1-q_0}\right) + \frac{1}{1-q_0}\right] + \int_0^\infty \rho e^{-\rho t}\left(\log\left(\frac{q_t}{1-q_t}\right) + \frac{1}{1-q_t}\right)dt$$

Thus, the function to be maximized in the program above can now be written as:

$$\int_0^\infty e^{-\rho t}\left(W^a q_t - \frac{cu^b}{v^a}\frac{q_t}{(1-q_t)} + \frac{cu^a}{v^a}\right)dt - \frac{c}{\beta v^a}\int_0^\infty e^{-\rho t}\frac{\overset{\bullet}{q}_t}{q_t(1-q_t)^2}dt$$
$$= \int_0^\infty e^{-\rho t}\left(W^a q_t - \frac{cu^b}{v^a}\frac{q_t}{(1-q_t)} + \frac{cu^a}{v^a}\right)dt$$
$$- \frac{c}{\beta v^a}\left[\int_0^\infty \rho e^{-\rho t}\left(\log\left(\frac{q_t}{1-q_t}\right) + \frac{1}{1-q_t}\right)dt\right]$$
$$+ \frac{c}{\beta v^a}\left[\log\left(\frac{q_0}{1-q_0}\right) + \frac{1}{1-q_0}\right]$$
$$= \int_0^\infty e^{-\rho t}\left\{W^a q_t - \left(\frac{\frac{cu^b}{v^a}q_t + \frac{c\rho}{\beta v^a}}{1-q_t}\right) - \frac{c\rho}{\beta v^a}\log\left(\frac{q_t}{1-q_t}\right)\right\}dt + K$$

where K is a constant ($K = \frac{c}{\beta v^a}\left[\log\left(\frac{q_0}{1-q_0}\right) + \frac{1}{1-q_0}\right] + \frac{cu^a}{\rho v^a}$). Our optimal control problem therefore collapses to

$$\max_{0 \le G_t \le \overline{G}} \int_0^\infty e^{-\rho t} R(q_t) dt$$

where

$$R(q_t) = W^a q_t - \left(\frac{\frac{cu^b}{v^a} q_t + \frac{c\rho}{\beta v^a}}{1-q_t}\right) - \frac{c\rho}{\beta v^a} \log\left(\frac{q_t}{1-q_t}\right)$$

Denote for simplicity $\beta = v^a = 1$, then

$$R(q_t) = W^a q_t - c\left(\frac{u^b q_t + \rho}{1-q_t}\right) - c\rho \log\left(\frac{q_t}{1-q_t}\right)$$

This completes the proof of this lemma. $\square$

Proof of Lemma 2. Observe that

$$R'(q) = W^a - c\frac{\rho + u^b q}{q(1-q)^2}$$

Now, consider any point q satisfying the first order condition

$$R'(q) = W^a - c\frac{\rho + u^b q}{q(1-q)^2} = 0$$

Note, first, that the function $\Theta(q) = c\frac{\rho + u^b q}{q(1-q)^2}$ is such that $\lim_{q \to 0} \Theta(q) = \lim_{q \to 1} \Theta(q) = +\infty$. Moreover, it is easy to see that

$$\Theta'(q) = c\frac{\left[2q\left(\rho + u^b q\right) - \rho(1-q)\right]}{q^2(1-q)^3}$$

and that there exists a threshold $\widetilde{q} \in (0, 1)$ such that the numerator is positive if and only if $q \ge \widetilde{q}$. Therefore the function $\Theta(q)$ is decreasing in q for $q \le \widetilde{q}$ and increasing in q for $q \ge \widetilde{q}$, reaching a minimum at $\widetilde{q}$. Hence the function $R'(q)$ is increasing in q for $q \le \widetilde{q}$ and decreasing in q for all $q \ge \widetilde{q}$, and $R'(1) = -\infty$. Define $\overline{W} = \Theta(\widetilde{q})$ and the lemma follows trivially from the expression of $R'(q)$.

Now when $W^a < \overline{W}$, $R'(q) < \Theta(\widetilde{q}) - \Theta(q) < 0$ for all $q \in [0, 1]$ and the function $R(q)$ is decreasing in q. When $W^a > \overline{W}$, then $\lim_{q \to 0} R'(q) = -\infty$, and $R'(\widetilde{q}) = W^a - \Theta(\widetilde{q}) > 0$. Thus, there exists a unique $q_L \in (0, \widetilde{q})$ such that $R'(q_L) = 0$ and $R''(q_L) > 0$ (a local minimum). At the same time, on the segment $[\widetilde{q}, 1]$, we also have that $R'(q)$ is decreasing in q and that $\lim_{q \to 1} R'(q) = -\infty$ and $R'(\widetilde{q}) > 0$. Thus, there exists a unique $q_H \in (\widetilde{q}, 1)$ such that $R'(q_H) = 0$ and $R''(q_H) < 0$ (a local maximum). The comparative statics are immediate once one observes that $R'(q)$ shifts up with W^a, that the function $\Theta(q)$ shifts up with c, ρ and u^b and that $R''(q_L) > 0$ and $R''(q_H) < 0$. Note also that q_L and q_H are not affected by changes in u^a (as the function $R(q_t)$ does not depend on u^a). $\square$

Proof of Lemma 3. Given that $R(q)$ is decreasing in q for $q \le \widetilde{q}$, with $\lim_0 R(q) = +\infty$ and $R(q_L) < R(q_H)$, then there exists a unique point $q_S^* \in (0, q_L)$ such that $R(q_S^*) = R(q_H)$. By having a generic parameter $\alpha \in \{W, c, \rho, u^b\}$, we have

$$\frac{\partial R(q_S^*)}{\partial \alpha} + \frac{dR(q_S^*)}{dq}\frac{dq_S^*}{d\alpha} = \frac{\partial R(q_H)}{\partial \alpha} + \frac{dR(q_H)}{dq}\frac{dq_H}{d\alpha} = \frac{\partial R(q_H)}{\partial \alpha}$$

Thus

$$\frac{dq_S^*}{d\alpha} = \frac{\frac{\partial R(q_H)}{\partial \alpha} - \frac{\partial R(q_S^*)}{\partial \alpha}}{\frac{dR(q_S^*)}{dq}}$$

Given that $\frac{dR(q_S^*)}{dq} < 0$ (as $q_S^* \in (0, q_L)$), the sign of $\frac{dq_S^*}{d\alpha}$ is the same as the sign of $\frac{\partial R(q_S^*)}{\partial \alpha} - \frac{\partial R(q_H)}{\partial \alpha}$. Now

$$\frac{\partial^2 R}{\partial \alpha \partial q} > 0 \text{ for } \alpha = W^a$$

$$\frac{\partial^2 R}{\partial \alpha \partial q} < 0 \text{ for } \alpha = c, \rho \text{ and } u^b$$

Hence given that $q_S^* < q_H$, we immediately get the result that q_S^* is a decreasing function of W^a, and an increasing function of c, ρ and u^b. $\square$

Proof of Proposition 1.

(i) When $W^a < \overline{W}$, the function $R(q)$ is decreasing for all $q \in (0, 1)$. Also, all trajectories $q(t, G^a(t))$ starting at $q(0) = \frac{u^a}{u^a+u^b}$ for $G^a(t) > 0$ is such that $q(t, G^a(t)) > q(0) = \frac{u^a}{u^a+u^b}$. Hence the value of that trajectory $V(q(., G^a))$ is such that

$$V(q(., G^a)) = \int_0^\infty e^{-\rho t} R(q(t, G^a(t)))dt < \int_0^\infty e^{-\rho t} R(q(0))dt = \frac{R(q(0))}{\rho}$$

and therefore is dominated by the trajectory that stays permanently at $q(0)$ with $G^a(t) = 0$ for all t.

(ii) Consider now the case where $W^a > \overline{W}$. From Lemma 2 and 3, we know that there exist two interior points q_L and $q_H \in (0, 1)$ such that $R(q)$ is decreasing in $q \in [0, q_L]$, increasing in $q \in [q_L, q_H]$ and decreasing in $q > q_H$. Moreover there exists a unique point $q_S^* \in (0, q_L)$ such that that $R(q_S^*) = R(q_H)$.

$(ii1)$ Consider first the case where $q(0) < q_S^*$. Then, we know that $V(q(0)) > V(q)$ for any $q > q(0)$. Therefore, given that any trajectory $q(., G^a)$ with $G^a(t) > 0$ and starting at $q(0)$ is such that $q(t, G^a(t)) > q(0)$, it follows as before that such a trajectory is dominated by staying permanently at $q(0)$ with $G^a(t) = 0$ for all t.

$(ii2)$ Consider now the case where $q(0) \in (q_S^*, q^F)$ where $q^F = \min\left[q_H, q(\overline{G}^a)\right]$ and the segment $\left[q(0), q(\overline{G}^a)\right]$ that contains q_F. We know that all points q of the segment $\left[q(0), q^F\right]$ can be reached by a feasible $q(t, G^a)$ trajectories starting at $q(0) < q(\overline{G}^a)$ and satisfying the cultural evolution equation (9) with $G^a(t) \in \left[0, \overline{G}^a\right]$. In the segment $\left[q(0), q(\overline{G}^a)\right]$, the function $V(q)$ has at most two local maximum points: $q = q(0)$ (when $q(0) < q_L$, see Figure 1) and $q = q^F$.

Following the result of Spence and Starrett (1975), it follows that the optimal trajectory is a Most Rapid Approach Path (MRAP) to one of these local maxima.

Whenever $q(0)$ is a local maximum (i.e. when $q(0) < q_L$), the MRAP starting from $q_0 = q(0)$ is trivially staying there permanently with $G^a(t) = 0$ for all t.

For $q = q^F$, the MRAP from the initial point $q_0 = q(0)$ is to have $G^a(t) = \overline{G}^a$ up to the point where the trajectory $q(t, G^a)$ reaches q^F and then to stay there indefinitely. When $q_H > q(\overline{G}^a)$, $q^F = q(\overline{G}^a)$ and the MRAP reaches q^F as a steady state so the control solution is to have maximum capacity at $G^a(t) = \overline{G}^a$ at all time. When conversely $q_H < q(\overline{G}^a)$, then $q^F = q_H$ and the MRAP reaches q_H in some finite time T. After that time, the optimal control solution is to have a "singular" interior arc policy $G^a(t) = G^F < \overline{G}^a$ that ensures that the system stays permanently at the steady state q_H. This is ensured by choosing G^F such that $\dot{q}(t) = 0$ at $q = q_H$ and $G = G^F$. This leads to:

$$(1 - q_H)\left(u^a + G^F\right) - q_H u^b = 0$$

or

$$G^F = -u^a + \frac{q_H u^b}{1 - q_H}$$

When $q(0) > q_L$, there is a unique local maximum in the segment $\left[q(0), q(\overline{G}^a)\right]$ and therefore the optimal trajectory is simply the MRAP leading to $q = q^F$.

When $q(0) < q_L$ one needs to compare the value of the cultural leader along the MRAP trajectories leading to the two local maxima to decide which one is optimal.

Consider first the case where $q_H < q(\overline{G}^a)$, and the MRAP from $q(0)$ to $q^F = q_H$ reaches q^F in some finite time $T\,(q(0)) = T_0$. The value along that path can be written as

$$V^F(q(0)) = \int_0^{T_0} e^{-\rho t} R(q^+(t, q(0)))dt + e^{-\rho T_0} \frac{R(q^F)}{\rho}$$

where $q^+(t, q(0))$ is the solution of the differential equation

$$\frac{dq_t}{dt} = q_t(1 - q_t)\left[(1 - q_t)\left[u^a + \overline{G}^a\right] - q_t u^b\right]$$
$$\text{and} \quad q_0 = q(0)$$
$$\text{and} \quad q^+(T_0, q(0)) = q^F$$

The value along the MRAP at $q(0)$ is given by $R(q(0))/\rho$. Define then the function

$$\Omega(q) = V^F(q) - \frac{R(q)}{\rho}$$

that is the difference between the value $V^F(q)$ of the MRAP starting from a value q to q^F and the value of staying permanently at that point q. Then we have

$$\Omega(q(0)) = V^F(q(0)) - \frac{R(q(0))}{\rho}$$

Obviously

$$\begin{aligned}\Omega(q_S^*) &= \int_0^{T(q_S^*)} e^{-\rho t}R(q^+(t,q_S^*))dt + e^{-\rho T(q_S^*)}\left[\frac{R(q^F)}{\rho}-\frac{R(q_S^*)}{\rho}\right]-\frac{1-e^{-\rho T(q_S^*)}}{\rho}R(q_S^*)\\ &< \frac{1-e^{-\rho T(q_S^*)}}{\rho}R(q_S^*) + e^{-\rho T(q_S^*)}\left[\frac{R(q^F)}{\rho}-\frac{R(q_S^*)}{\rho}\right]-\frac{1-e^{-\rho T(q_S^*)}}{\rho}R(q_S^*)\\ &= \frac{e^{-\rho T(q_S^*)}}{\rho}\left[R(q^F)-R(q_S^*)\right]<0\end{aligned}$$

as $R(q_S^*) > R(q^+(t,q_S^*))$ and $R(q_S^*) > R(q)$ for $q \in \left(q_S^*, q^F\right)$ while

$$\begin{aligned}\Omega(q_L) &= \int_0^{T(q_L)} e^{-\rho t}R(q^+(t,q_L))dt + e^{-\rho T(q_L)}\left[\frac{R(q^F)}{\rho}-\frac{R(q_L)}{\rho}\right]-\frac{1-e^{-\rho T(q_L)}}{\rho}R(q_L)\\ &> \frac{1-e^{-\rho T(q_L)}}{\rho}R(q_L) + e^{-\rho T(q_L)}\left[\frac{R(q^F)}{\rho}-\frac{R(q_L)}{\rho}\right]-\frac{1-e^{-\rho T(q_L)}}{\rho}R(q_L)\\ &> \frac{e^{-\rho T(q_L)}}{\rho}[R(q^F)-R(q_L)]>0\end{aligned}$$

as $q^+(t,q_L) > q_L$ and $R(.)$ is increasing in $q \in [q_L, q_H]$.

Now for $q \in \left(q_S^*, q_L\right)$, and h small enough that $q+h \in \left(q_S^*, q_L\right)$, denote $T(h)$ the time necessary to go from point q to point $q+h$ along the trajectory solution $q^+(t,.)$ starting at q (i.e. $q^+(T(h),q) = q+h$). Then we can write

$$V^F(q) = \int_0^{T(h)} e^{-\rho t}R(q^+(t,q))dt + e^{-\rho T(h)}V^F(q+h)$$

For h small enough, $\left(q^+(T(h),q)-q\right)/T(h) \simeq \dot{q}$. Thus $T(h) \simeq \frac{h}{f(q)}$ where $f(q) = q(1-q)\left[(1-q)\left[u^a+\overline{G}^a\right]-qu^b\right]$ and

$$V^F(q) \simeq \int_0^{\frac{h}{f(q)}} e^{-\rho t}R(q^+(t,q))dt + e^{-\rho\frac{h}{f(q)}}V^F(q+h)$$

or

$$\begin{aligned}V^F(q) &\simeq \left[R(q)+hR'(q)\right]\frac{(1-e^{-\rho\frac{h}{f(q)}})}{\rho} + e^{-\rho\frac{h}{f(q)}}(V^F(q)+V'^F(q)h)\\ &\simeq \left[R(q)+hR'(q)\right]\frac{h}{f(q)} + (1-\rho\frac{h}{f(q)})(V^F(q)+V'^F(q)h)\end{aligned}$$

As a consequence, taking the limit $h \to 0$,

$$V'^F(q) \simeq -\frac{R(q)}{f(q)} + \frac{\rho}{f(q)}V^F(q)$$

and

$$\Omega'(q) = V^F(q) - \frac{R(q)}{\rho}$$
$$= -\frac{R(q)}{f(q)} + \frac{\rho}{f(q)} V^F(q) - \frac{R'(q)}{\rho}$$

This implies that for any point $\widehat{q}^*$ such that $V^F(\widehat{q}^*) - \frac{R(\widehat{q}^*)}{\rho} = 0$, one has $\Omega'(\widehat{q}^*) = -\frac{R'(\widehat{q}^*)}{\rho} > 0$. Hence the uniqueness of $\widehat{q}^*$.

The preceding discussion implies that there exists a unique $\widehat{q}^* \in \left(q_S^*, q_L\right)$ such that $\Omega(\widehat{q}^*) = 0$ and that $\Omega(q(0)) > 0$ if and only if $q(0) > \widehat{q}^*$. It follows that if $q(0) < \widehat{q}^*$, the optimal trajectory for the leader is to stay permanently at $q(0)$ with $G^a(t) = 0$, while for $q(0) > \widehat{q}^*$, the optimal trajectory for the leader is the MRAP from $q(0)$ to q_H with $G^a(t) = \overline{G}^a$ until q_H is reached in finite time $T_0 = T(q(0))$, and then to stay permanently there with $G^a(t) = G^{a*} < \overline{G}^a$.

Consider now the case where $q_H > q(\overline{G}^a)$, and the MRAP from $q(0)$ to $q^F = q(\overline{G}^a)$ reaches q^F for $T = +\infty$. The same argument as before can be performed for $T = +\infty$ and the same result as before applies.

$(ii3)$ When $q(0) \geq q_H$, along all possible trajectory $q(t, q(0))$ for $G^a(t) \in \left[0, \overline{G}^a\right]$ one has $R(q(t, q(0))) < R(q(0))$ as the function $R(.)$ is decreasing in $q \in \left[q_H, q(\overline{G}^a)\right]$. Hence the optimal trajectory is for the leader to remain inactive at all time (i.e. $G^a(t) = 0$) and the cultural state stays at $q(0)$. □

Proof of Proposition 2. When $q(0) \in (\widehat{q}^*, q_H)$, $\widehat{q}^*$ is given by the following equation:

$$\Omega\left(\widehat{q}^*, \alpha\right) = \int_0^{T(\widehat{q}^*, \alpha)} e^{-\rho t} R(q^+(t, \widehat{q}^*), \alpha) dt + e^{-\rho T(\widehat{q}^*, \alpha)} \frac{R(q^F, \alpha)}{\rho} - \frac{R(\widehat{q}^*, \alpha)}{\rho} = 0 \tag{29}$$

where α is a generic parameter and $T(\widehat{q}^*, \alpha)$ is such that $q^+(T(\widehat{q}^*, \alpha), \widehat{q}^*)) = q^F$.

(i) For $\alpha = W, c$, we have:

$$\frac{\partial \Omega}{\partial \alpha} = \int_0^{T(\widehat{q}^*)} e^{-\rho t} \frac{\partial R}{\partial \alpha}(q^+(t, \widehat{q}^*), \alpha) dt + \frac{e^{-\rho T(\widehat{q}^*)}}{\rho} \left[R'(q^F) \frac{\partial q^F}{\partial \alpha} + \frac{\partial R(q^F)}{\partial \alpha} \right] - \frac{1}{\rho} \frac{\partial R(\widehat{q}^*)}{\partial \alpha}$$

Notice that $R(q^+(T, \widehat{q}^*), \alpha) = R(q^F, \alpha)$ and therefore the term in $\frac{\partial T}{\partial \alpha}$ disappears. Moreover,

$$\frac{e^{-\rho T(\widehat{q}^*)}}{\rho} \left[R'(q^F) \frac{\partial q^F}{\partial \alpha} + \frac{\partial R(q^F)}{\partial \alpha} \right]$$
$$= \begin{cases} \frac{e^{-\rho T(\widehat{q}^*)}}{\rho} \frac{\partial R(q^F)}{\partial \alpha} & \text{when } q^F = q_H \quad \text{and } T(\widehat{q}^*) < \infty \\ 0 & \text{when } q^F = q(\overline{G}^a) \quad \text{and } T(\widehat{q}^*) = \infty \end{cases}$$

When $q^F = q_H < q(\overline{G}^a)$, we have:

$$\frac{\partial \Omega}{\partial \alpha} = \int_0^{T(\widehat{q}^*)} e^{-\rho t} \frac{\partial R}{\partial \alpha}(q^+(t, \widehat{q}^*), \alpha) dt + \frac{e^{-\rho T(\widehat{q}^*)}}{\rho} \left[\frac{\partial R(q^F, \alpha)}{\partial \alpha} \right] - \frac{1}{\rho} \frac{\partial R(\widehat{q}^*)}{\partial \alpha}$$

$$= \int_0^{T(\widehat{q}^*)} e^{-\rho t}\left[\frac{\partial R}{\partial\alpha}(q^+(t,\widehat{q}^*),\alpha) - \frac{\partial R(\widehat{q}^*,\alpha)}{\partial\alpha}\right]dt$$
$$+ \int_{T(\widehat{q}^*)}^{\infty} e^{-\rho t}\left[\frac{\partial R(q^F,\alpha)}{\partial\alpha} - \frac{\partial R(\widehat{q}^*,\alpha)}{\partial\alpha}\right]dt$$

as $R(q,W,\rho,c,u^b) = Wq - c\frac{u^b q+\rho}{1-q} - c\rho\log\left(\frac{q}{1-q}\right)$,

$$\frac{\partial^2 R}{\partial\alpha\partial q} > 0 \text{ for } \alpha = W \text{ and } \frac{\partial^2 R}{\partial\alpha\partial q} < 0 \text{ for } \alpha = c$$

Hence given that $q^+(t,\widehat{q}^*) > \widehat{q}^*$, it follows that $\frac{\partial R(q^+,\alpha)}{\partial\alpha} - \frac{\partial R(\widehat{q}^*)}{\partial\alpha}$ is positive for $\alpha = W$ and negative for $\alpha = c$. Hence observing that $q^+(T,\widehat{q}^*) = q_H$, it follows that $\frac{\partial\Omega}{\partial\alpha}$ is positive for $\alpha = W$ and negative for $\alpha = c$. Given that $\Omega_q(\widehat{q}^*,\alpha) > 0$ and

$$\frac{\partial\widehat{q}^*}{\partial\alpha} = -\frac{\Omega_\alpha(\widehat{q}^*,\alpha)}{\Omega_q(\widehat{q}^*,\alpha)}$$

it follows that

$$\frac{\partial\widehat{q}^*}{\partial\alpha} < 0 \text{ for } \alpha = W \text{ and } \frac{\partial\widehat{q}^*}{\partial\alpha} > 0 \text{ for } \alpha = c$$

When $q^F = q(\overline{G}^a) < q^H$

$$\frac{\partial\Omega}{\partial\alpha} = \int_0^{\infty} e^{-\rho t}\frac{\partial R}{\partial\alpha}(q^+(t,\widehat{q}^*),\alpha)dt - \frac{1}{\rho}\frac{\partial R(\widehat{q}^*)}{\partial\alpha}$$

and similarly

$$\frac{\partial\Omega}{\partial\alpha} = \int_0^{\infty} e^{-\rho t}\left[\frac{\partial R}{\partial\alpha}(q^+(t,\widehat{q}^*),\alpha) - \frac{\partial R(\widehat{q}^*,\alpha)}{\partial\alpha}\right]dt$$

A similar argument as above provides also $\frac{\partial\widehat{q}^*}{\partial\alpha} < 0$ for $\alpha = W$ and $\frac{\partial\widehat{q}^*}{\partial\alpha} > 0$ for $\alpha = c$.

(ii) Consider now $\alpha = \rho$ and consider the case when $q^F = q_H < q(\overline{G}^a)$ (a similar argument applies to the other case $q^F = q(\overline{G}^a) < q_H$). Then

$$\frac{\partial\Omega}{\partial\rho} = \int_0^{T(\widehat{q}^*)} e^{-\rho t}\left[\frac{\partial R(q^+)}{\partial\rho} - \frac{\partial R(\widehat{q}^*)}{\partial\rho}\right]dt + \int_{T(\widehat{q}^*)}^{\infty} e^{-\rho t}\left[\frac{\partial R(q_H)}{\partial\rho} - \frac{\partial R(\widehat{q}^*)}{\partial\rho}\right]dt$$
$$-\left[\int_0^{T(\widehat{q}^*)} te^{-\rho t}\left[R(q^+) - R(\widehat{q}^*)\right]dt + \int_{T(\widehat{q}^*)}^{\infty} te^{-\rho t}\left[R(q_H) - R(\widehat{q}^*)\right]dt\right] \tag{30}$$

where (with abuse of notation) we write $q^+ = q^+(t,\widehat{q}^*)$. Now

$$\frac{\partial^2 R}{\partial\rho\partial q} < 0$$

Hence given that $q^+ = q^+(t, \widehat{q}^*) > \widehat{q}^*$ and that $q_H > \widehat{q}^*$, the first line of the previous expression has a negative sign. Denote the expression of the bracket of the second row as

$$I = \int_0^{T(\widehat{q}^*)} te^{-\rho t}\left[R(q^+) - R(\widehat{q}^*)\right]dt + \int_{T(\widehat{q}^*)}^{\infty} te^{-\rho t}\left[R(q_H) - R(\widehat{q}^*)\right]dt$$

Also denote the function

$$V(t) = \int_0^t e^{-\rho u}\left[R(q^+(u, \widehat{q}^*)) - R(\widehat{q}^*)\right]du$$

Then integration by part gives

$$\begin{aligned} I &= [tV(t)]_0^{T(\widehat{q}^*)} - \int_0^{T(\widehat{q}^*)} V(t)dt + \int_{T(\widehat{q}^*)}^{\infty} te^{-\rho t}\left[R(q_H) - R(\widehat{q}^*)\right]dt \\ &= T(\widehat{q}^*)V(T(\widehat{q}^*)) - \int_0^{T(\widehat{q}^*)} V(t)dt + \int_{T(\widehat{q}^*)}^{\infty} te^{-\rho t}\left[R(q_H) - R(\widehat{q}^*)\right]dt \\ &= T(\widehat{q}^*)V(T(\widehat{q}^*)) - \int_0^{T(\widehat{q}^*)} V(t)dt + \left[\frac{e^{-\rho T(\widehat{q}^*)}}{\rho^2} + \frac{T(\widehat{q}^*)e^{-\rho T(\widehat{q}^*)}}{\rho}\right]\left[R(q_H) - R(\widehat{q}^*)\right] \end{aligned}$$

Note that we also have (by the definition of $\widehat{q}^*$)

$$\int_0^{T(\widehat{q}^*)} e^{-\rho t}\left[R(q^+) - R(\widehat{q}^*)\right]dt + \int_{T(\widehat{q}^*)}^{\infty} e^{-\rho t}\left[R(q_H) - R(\widehat{q}^*)\right]dt = 0$$

or

$$V(T(\widehat{q}^*)) = -\frac{e^{-\rho T(\widehat{q}^*)}}{\rho}\left[R(q_H) - R(\widehat{q}^*)\right]$$

Now denote the function $\Phi(T)$ as

$$\Phi(T) = TV(T) - \int_0^T V(t)dt + \left[\frac{e^{-\rho T}}{\rho^2} + \frac{Te^{-\rho T}}{\rho}\right]\left[R(q_H) - R(\widehat{q}^*)\right]$$

Then

$$\begin{aligned} I &= \Phi(T(\widehat{q}^*)) \\ &= T(\widehat{q}^*)V(T(\widehat{q}^*)) - \int_0^{T(\widehat{q}^*)} V(t)dt + \left[\frac{e^{-\rho T(\widehat{q}^*)}}{\rho^2} + \frac{T(\widehat{q}^*)e^{-\rho T(\widehat{q}^*)}}{\rho}\right]\left[R(q_H) - R(\widehat{q}^*)\right] \end{aligned}$$

or

$$I = -\int_0^{T(\widehat{q}^*)} V(t)dt + \frac{e^{-\rho T(\widehat{q}^*)}}{\rho^2} \cdot \left[R(q_H) - R(\widehat{q}^*)\right] \tag{31}$$

Now notice that

$$V(t) < \left[R(q_H) - R(\widehat{q}^*)\right] \cdot \frac{1 - e^{-\rho t}}{\rho}$$

and

$$\int_0^{T(\widehat{q}^*)} V(t)dt < \left[R(q_H) - R(\widehat{q}^*)\right] \cdot \int_0^{T(\widehat{q}^*)} \frac{1 - e^{-\rho t}}{\rho} dt$$
$$= \left[R(q_H) - R(\widehat{q}^*)\right] \cdot \left[\frac{T(\widehat{q}^*)}{\rho} + \frac{e^{-\rho T(\widehat{q}^*)}}{\rho^2} - \frac{1}{\rho^2}\right]$$

Thus using (31), it follows that

$$I > \frac{\left[R(q_H) - R(\widehat{q}^*)\right]}{\rho^2} \left[1 - T(\widehat{q}^*)\rho\right]$$

and $I > 0$ when $\rho < 1/T(\widehat{q}^*)$. Therefore the second row of (30) is negative when $\rho < 1/T(\widehat{q}^*)$ and we obtain that

$$\frac{\partial \Omega}{\partial \rho} < 0 \quad \text{for } \rho \text{ small enough}$$

As a result

$$\frac{\partial \widehat{q}^*}{\partial \rho} > 0 \quad \text{for } \rho \text{ small enough}$$

Note that the comparative statics for u^b cannot be easily signed. Indeed because $\frac{\partial^2 R}{\partial u^b \partial q} < 0$, a MRAP towards the lower value q^F is less attractive to the cultural leader and as a consequence, the leader is less likely to be active when u^b increases. At the same time though, an increase in u^b affects also the speed of cultural change. Indeed, along a MRAP towards the steady state value q^F, the speed of convergence is reduced. This initially has a positive value effect for the MRAP towards q^F in the time interval $\left[0, T^L(\widehat{q}^*)\right]$ with $T^L(\widehat{q}^*)$ given by $R(q^+(t, \widehat{q}^*) = R(q_L)$ in which the cultural value $R(q^+(t, \widehat{q}^*))$ is smaller than $R(\widehat{q}^*)$. However, it has a negative effect in the time interval where $R(q^+(t, \widehat{q}^*))$ becomes larger than $R(\widehat{q}^*)$ and approaches q^F at a lower speed. The overall effect of an increase of u^b on the threshold $\widehat{q}^*$ is therefore ambiguous. □

Proof of Proposition 3. The only non trivial part (all the other comparative statics results are relatively straightforward) is to consider the effect of the other group incentives (i.e. u^b) on the leader effort. Simple differentiation provides that

$$\frac{\partial q_H}{\partial u^b} = -\frac{R'_{qu^b}}{R'_{qq}} = -\frac{(q_H)^2 (1 - q_H)}{\left[2q_H \left(\rho + u^b q_H\right) - \rho(1 - q_H)\right]}$$

As

$$G^F = -u^a + \frac{q_H u^b}{1-q_H}$$

we get

$$\frac{\partial G^F}{\partial u^b} = \frac{\partial q_H}{\partial u^b}\frac{u^b}{(1-q_H)^2} + \frac{q_H}{1-q_H} \tag{32}$$
$$= -\frac{(q_H)^2}{\left[2q_H\left(\rho+u^b q_H\right)-\rho(1-q_H)\right]}\frac{u^b}{(1-q_H)} + \frac{q_H}{1-q_H}$$

Clearly the first term of the RHS of (32) is negative while the second term is positive. In general the sign of $\partial G^F/\partial u^b$ is ambiguous. However when $u^b \simeq 0$, the first term vanishes and remains only the second positive term $q_H/(1-q_H)$. Therefore $\left(\frac{\partial G^F}{\partial u^b}\right)_{u^b=0} > 0$.

Furthermore $\left(\partial G^F/\partial u^b\right) < 0$ when

$$1 - \frac{u^b}{\left[2\left(\rho+u^b q_H\right)-\rho(\frac{1-q_H}{q_H})\right]} < 0$$

Given that the function $\Lambda(q) = \frac{u^b}{\left[2\left(\rho+u^b q\right)-\rho(\frac{1-q}{q})\right]}$ is decreasing in q and that $q_H < q\left(\overline{G}^a\right)$ (for $G^F < \overline{G}^a$) one has

$$1 - \frac{u^b}{\left[2\left(\rho+u^b q_H\right)-\rho(\frac{1-q_H}{q_H})\right]} < 1 - \frac{u^b}{\left[2\left(\rho+u^b q\left(\overline{G}^a\right)\right)-\rho(\frac{1-q\left(\overline{G}^a\right)}{q\left(\overline{G}^a\right)})\right]}$$
$$= 1 - \frac{u^b}{\left[2\left(\rho+\frac{u^b\left(u^a+\overline{G}^a\right)}{u^a+\overline{G}^a+u^b}\right)-\rho(\frac{u^b}{u^a+\overline{G}^a})\right]}$$

Now $1 - \frac{u^b}{\left[2\left(\rho+\frac{u^b\left(u^a+\overline{G}^a\right)}{u^a+\overline{G}^a+u^b}\right)-\rho(\frac{u^b}{u^a+\overline{G}^a})\right]} < 0$ when $2\left(\rho+\frac{\left(u^a+\overline{G}^a\right)u^b}{u^a+\overline{G}^a+u^b}\right) < u^b\left[\frac{\rho}{u^a+\overline{G}^a}+1\right]$. Now the function $Q(u) = 2\left(\rho+\frac{u\left(u^a+\overline{G}^a\right)}{u+u^a+\overline{G}^a}\right)$ is increasing concave in u with $Q(0) = 2\rho$ and $\lim_{u\to\infty} Q(u) = 2\left(\rho+u^a+\overline{G}^a\right)$ while the function $S(u) = u\left[\frac{\rho}{u^a+\overline{G}^a}+1\right]$ is a linear increasing function of u with $S(0) = 0$ and $\lim_{u\to\infty} S(u) = +\infty$. Hence by continuity there is a unique $\overline{u} > 0$ such that $Q(u^b) < S(u^b)$ if and only if $u^b > \overline{u}$. From this it follows that when $u^b > \overline{u}$ one has $1 - \frac{u^b}{\left[2\left(\rho+\frac{u^b\left(u^a+\overline{G}^a\right)}{u^a+\overline{G}^a+u^b}\right)-\rho(\frac{u^b}{u^a+\overline{G}^a})\right]} < 0$ and therefore $1 - \frac{u^b}{\left[2\left(\rho+u^b q_H\right)-\rho(\frac{1-q_H}{q_H})\right]} < 0$, implying that $\left(\partial G^F/\partial u^b\right) < 0$ for u^b large enough. □

Proof of Proposition 4. Let us start by the following two lemmas.

Lemma 5. *Assume that (26) holds and that both leaders have the same discount rate $\rho^a = \rho^b = \rho$. Then the open-loop best response of leader i to leader $j \neq i$ choosing permanently $G^j = \overline{G}$ is $G^i = \overline{G}$.*

Proof of Lemma 5. Consider the case of leader a, as the case of leader b can be treated in a symmetric way. When leader b is choosing permanently $G^b = \overline{G}$, the choice for leader a is simply between playing $G^a = \overline{G}$ at all time and staying permanently at the initial cultural dynamics $q(0) = \frac{1}{2}$ or choosing some policy path $\left(G_t^a\right)_t$ such that $G_t^a < \overline{G}$ at least during some time interval. It is immediate to see that the cultural trajectory $\left(q_t(G^a, \overline{G})\right)_{t\geq 0}$ implied by such policy path $\left(q_t(G^a, \overline{G})\right)_{t\geq 0}$ is such that

$$\frac{u}{2u+\overline{G}} \leq q_t(G^a, \overline{G}) < q(0) = \frac{1}{2}$$

since $\frac{u}{2u+\overline{G}}$ is the steady state of the decreasing monotonic cultural dynamics with initial condition $q(0) = \frac{1}{2}$ when $G^b = \overline{G}$ and $G^a = 0$ at all time. Observing that the function $R\left(q_t, u+\overline{G}, \rho, \frac{W}{c}\right)$ is increasing in the interval $q_t \in \left[q_L\left(W/c, \rho, u+\overline{G}\right), \frac{1}{2}\right]$ and that $q_t(G^a, \overline{G}) \in \left[\frac{u}{2u+\overline{G}}, \frac{1}{2}\right[$ since $q_L\left(W/c, \rho, u+\overline{G}\right) < \frac{u}{2u+\overline{G}}$ (see Fig. 5), we then have that, along such policy path, the value $V^a(\left(q_t(G^a, \overline{G})\right)_{t\geq 0})$ for leader a is such that

$$V^a\left(\left(q_t(G^a, \overline{G})\right)_{t\geq 0}\right) = \int_0^\infty e^{-\rho t} R\left(q_t(G^a, \overline{G}), u+\overline{G}, \rho, \frac{W}{c}\right) dt + 2c + \frac{u}{\rho}$$

$$< \int_0^\infty e^{-\rho t} R\left(\frac{1}{2}, u+\overline{G}, \rho, \frac{W}{c}\right) dt + 2c + \frac{u}{\rho} = \frac{R\left(\frac{1}{2}, u+\overline{G}, \rho, \frac{W}{c}\right)}{\rho} + 2c + \frac{u}{\rho}$$

The last term is the value for leader a to choose $G^a = \overline{G}$ at all time and therefore stay permanently at the initial cultural dynamics $q(0) = \frac{1}{2}$. It follows that the open-loop best response of leader a to leader B choosing permanently $G^b = \overline{G}$ is also $G^a = \overline{G}$. $\square$

Lemma 6. *Assume that (26) and (27) hold. Then, for each leader $i \in \{a, b\}$ playing $G^i = \overline{G}$ is a dominant strategy.*

Proof of Lemma 6. Under the more restricted assumption (27), let us show that playing $G^i = \overline{G}$ is a dominant strategy of any leader i to the other leader j playing permanently a constant value $G^j = G \in \left[0, \overline{G}\right]$. Indeed, take leader i. His payoff function is $R\left(q_t, u+G, \rho, \frac{W}{c}\right)$. Given that $q_L\left(W/c, \rho, u+G\right)$ is increasing in G and that $q_H\left(W/c, \rho, u+G\right)$ is decreasing in G, we have for $G \leq \overline{G}$

$$q_L\left(W/c, \rho, u+G\right) < \frac{u}{2u+\overline{G}} \leq \frac{u}{2u+G} < \frac{1}{2}$$

$$\frac{1}{2} \leq \frac{u+\overline{G}}{2u+G+\overline{G}} < \frac{u+\overline{G}}{2u+\overline{G}} < q_H\left(W/c, \rho, u+G\right)$$

Note that for any policy profile $G_t^a \in \left[0, \overline{G}\right]$, the cultural dynamics starting at $q(0) = 1/2$ stays within the interval $\left[\frac{u}{2u+G}, \frac{u+\overline{G}}{2u+G+\overline{G}}\right]$, which is included in the interval

$$[q_L(W/c, \rho, u+G), q_H(W/c, \rho, u+G)]$$

for which the function $R\left(q_t, u+G, \rho, \frac{W}{c}\right)$ is increasing in q. From this, it follows that the optimal trajectory for leader a is to play the MRAP going from $q(0) = 1/2$ to the steady state $\frac{u+\overline{G}}{2u+G+\overline{G}}$ with $G_t^a = \overline{G}$ at all time t. Hence the open loop best response of leader a to leader b playing permanently $G_t^b = G \in \left[0, \overline{G}\right]$ is to play $G_t^a = \overline{G}$ at all time t. $G_t^a = \overline{G}$ is thus a dominant open loop best response to any permanently constant effort by leader b. One obtains the same result by symmetry for leader b. □

From Lemmas 5 and 6, we conclude that under (27), $G_t^{a*} = G_t^{b*} = \overline{G}$ is the unique open loop Nash equilibrium such that both leaders choose constant public good provision. As a result, the two cultural leaders will play $\left(G^a, G^b\right) = (\overline{G}, \overline{G})$ at all time with the cultural dynamics remaining indefinitively at $q_t = q(0) = 1/2$. Finally, to obtain d_t^{a*} and d_t^{b*}, the socialization efforts of the parents, it suffices to plug $q_t = 1/2$ and $G^a = G^b = \overline{G}$ into (23) and (24). □

Proof of Proposition 5. When condition (28) holds, then, when $G_t^b = 0$, the choice for leader a is simply between playing $G^a = 0$ and staying permanently at the initial cultural dynamics $q(0) = \frac{1}{2}$ or choosing the MRAP that drives the cultural dynamics to the long-run steady state

$$q(G^a, \overline{G}) = \frac{u+\overline{G}}{2u+\overline{G}} > \frac{1}{2}$$

Given that $q(G^a, \overline{G}) < q_H(W/c, \rho^a, u)$, this MRAP is obtained by choosing permanently $G^a = \overline{G}$.

Consider now the problem of leader b when leader a chooses $G^a = \overline{G}$ at all time t. Given that now $W/c < \overline{W}\left(\rho^b, u+\overline{G}\right)$, the shape of the function $R(q, u+\overline{G}, \rho^b, \frac{W}{c})$ for the maximization problem of leader B is such that $R(q, u+\overline{G}, \rho^b, \frac{W}{c})$ is decreasing in q for all $q \in [0, 1]$ as depicted in Fig. 6b. Recall, as well, that, given that $G_t^b \in \left[0, \overline{G}\right]$ and

$$q(0) = \frac{1}{2} \in \left]\frac{u}{2u+\overline{G}}, \frac{u+\overline{G}}{2u+\overline{G}}, \right[$$

the cultural dynamics of q_t is given by:

$$\overset{\bullet}{\underline{q}}_t = \beta \underline{q}_t (1-\underline{q}_t)\left[\left(1-\underline{q}_t\right)(u+G_t^b) - \underline{q}_t\left(u+\overline{G}\right)\right]$$

and remains inside the interval $\left[\frac{u}{2u+\overline{G}}, \frac{1}{2}\right]$. As a result, the possible steady states of such dynamics also belong to this same interval. Given that $R(\underline{q}, u+\overline{G}, \rho^b, \frac{W}{c})$ is decreasing in q for all $\underline{q} \in \left[\frac{u}{2u+\overline{G}}, \frac{1}{2}\right]$, the optimal choice for leader b is then to choose the MRAP that makes the cultural dynamics $\underline{q}_t = 1 - q_t$ converging as fast as possible from $\underline{q}(0) = \frac{1}{2}$ to the lowest admissible steady state $\underline{q} = \frac{u}{2u+\overline{G}}$. This implies that it is optimal for leader b to play $G_t^b = 0$ at all time t.

From this, we see that, under condition (28), it is optimal for leader a to choose the MRAP $G^a = \overline{G}$ as the open-loop best response to leader b choosing permanently $G^b = 0$. Thus, the strategies $G_t^a = \overline{G}$ and $G_t^b = 0$ for all $t \geq 0$ is an open-loop Nash-equilibrium of the leaders' socialization game.

As for Proposition 4, it is again easy to see that, under the more restrictive condition (27), i.e.

$$q_L\left(W/c, \rho^a, u+\overline{G}\right) < \frac{u}{2u+\overline{G}} < \frac{1}{2} < \frac{u+\overline{G}}{2u+\overline{G}} < q_H\left(W/c, \rho^a, u+\overline{G}\right)$$

playing $G^a = \overline{G}$ is a dominant strategy for leader a to the other leader b playing permanently any constant value $G^b \in \left[0, \overline{G}\right]$. In such a case, and under (28), $G_t^{a*} = \overline{G}$ and $G_t^{b*} = 0$, $\forall t \geq 0$ is then the unique the open loop Nash equilibrium such that both leaders choose constant values of the public good. Finally, to obtain d_t^{a*} and d_t^{b*}, the socialization efforts of the parents, it suffices to plug $q_t = \frac{u+\overline{G}}{2u+\overline{G}}$ and $G^a = \overline{G}$ and $G^b = 0$ into (23) and (24). □

Appendix B. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jet.2018.01.019.

References

Ainsworth-Darnell, J.W., Downey, D.B., 1998. Assessing the oppositional culture explanation for racial/ethnic differences in school performance. Am. Sociol. Rev. 63, 536–553.

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

Alba, R.D., 1990. Ethnic Identity: The Transformation of White America. Yale University Press, New Haven.

Alesina, A., Reich, B., 2015. Nation Building. NBER Working Paper No. 18839.

Almagro, M., Andres, D., 2017. The construction of national identities. New York University. Unpublished manuscript.

Austen-Smith, D., Fryer Jr., R.D., 2005. An economic analysis of 'acting white'. Q. J. Econ. 120, 551–583.

Battu, H., Zenou, Y., 2010. Oppositional identities and employment for ethnic minorities. Evidence for England. Econ. J. 120, 52–71.

Berman, E., 2000. Sect, subsidy, and sacrifice: an economist's view of ultra-orthodox jews. Q. J. Econ. 115, 905–953.

Bernal, M.E., Knight, G.P., 1993. Ethnic Identity: Formation and Transmission among Hispanics and Other Minorities. State University of New York Press, Albany.

Bisin, A., Patacchini, E., Verdier, T., Zenou, Y., 2008. Are Muslim immigrants different in terms of cultural integration? J. Eur. Econ. Assoc. 6, 445–456.

Bisin, A., Patacchini, E., Verdier, T., Zenou, Y., 2011. Formation and persistence of oppositional identities. Eur. Econ. Rev. 55, 1046–1071.

Bisin, A., Patacchini, E., Verdier, T., Zenou, Y., 2016. Bend it like Beckham. Ethnic identity and integration. Eur. Econ. Rev. 90, 146–164.

Bisin, A., Verdier, T., 2000. Beyond the melting pot: cultural transmission, marriage, and the evolution of ethnic and religious traits. Q. J. Econ. 115, 955–988.

Bisin, A., Verdier, T., 2001. The economics of cultural transmission and the dynamics of preferences. J. Econ. Theory 97, 298–319.

Bisin, A., Verdier, T., 2011. The economics of cultural transmission and socialization. In: Benhabib, J., Bisin, A., Jackson, M.O. (Eds.), Handbook of Social Economics. Elsevier Science, Amsterdam, pp. 339–416.

Boyd, R., Richerson, P., 1985. Culture and the Evolutionary Process. University of Chicago Press, Chicago.

Carvalho, J.-P., 2018. Sacrifice and sorting in clubs. Forum Soc. Econ. https://doi.org/10.1080/07360932.2015.1125383. Forthcoming.

Carvalho, J.-P., Koyama, M., 2016. Jewish emancipation and schism: economic development and religious change. J. Comp. Econ. 44, 562–584.

Cavalli-Sforza, L., Feldman, M., 1981. Cultural Transmission and Evolution. A Quantitative Approach. Princeton University Press, Princeton.

Constant, A., Zimmermann, K.F., 2008. Measuring ethnic identity and its impact on economic behavior. J. Eur. Econ. Assoc. 6, 424–433.

Fordham, S., Ogbu, J.U., 1986. Black students' school success: coping with the burden of acting white. Urban Rev. 18, 176–206.

Fryer Jr., R.G., Torelli, P., 2010. An empirical analysis of 'acting white'. J. Public Econ. 94, 380–396.

Hauk, E., Mueller, H., 2015. Cultural leaders and the clash of civilizations. J. Confl. Resolut. 59, 367–400.

Iannaccone, L., 1992. Sacrifice and stigma: reducing free-riding in cults, communes, and other collectives. J. Polit. Econ. 100, 271–297.

Nielsen, R., 2012. Jihadi Radicalization of Muslim Clerics. Department of Government, Harvard University. Unpublished manuscript.

Nteta, T.M., Wallsten, K.J., 2012. Preaching to the choir? Religious leaders and American opinion on immigration reform. Soc. Sci. Q. 93, 891–910.

Panebianco, F., 2014. Socialization networks and the transmission of interethnic attitudes. J. Econ. Theory 150, 583–610.

Patacchini, E., Zenou, Y., 2016. Racial identity and education. Soc. Netw. 44, 85–94.

Phinney, J.S., 1990. Ethnic identity in adolescents and adults: review of research. Psychol. Bull. 108, 499–514.

Prummer, A., Siedlarek, J.-P., 2017. Community leaders and the preservation of cultural traits. J. Econ. Theory 168, 143–176.

Spence, M., Starrett, D., 1975. Most rapid approach paths in accumulation problems. Int. Econ. Rev. 16, 388–403.

Verdier, T., Zenou, Y., 2015. The role of cultural leaders in the transmission of preferences. Econ. Lett. 136, 158–161.

Yükleyen, A., Yurdakul, G., 2011. Islamic activism and immigrant integration: Turkish organizations in Germany. In: Immigrants and Minorities. In: Historical Studies in Ethnicity, Migration and Diaspora, vol. 29, pp. 64–85.

Yurdakul, G., Yükleyen, A., 2009. Islam, conflict, and integration: Turkish religious associations in Germany. Turkish Stud. 10, 217–231.

Cultural Leader and the Dynamics of Assimilation: Online Appendix

By Thierry Verdier[1] and Yves Zenou[2]

A Online Appendix A: An alternative formulation of the payoff structure

Instead of the utilities (5), we can have an alternative (not normalized) formulation of the payoffs structure. They can be written as follows:

$$\begin{aligned} U^a(A, G^a) &= (1-\delta)u^m + a + v^a G^a \\ U^a(B, G^a) &= u^m \\ U^b(B, G^b) &= u^m + b + v^b G^b \\ U^b(A, G^b) &= (1-\delta)u^m \end{aligned} \tag{A.1}$$

with parameters $u^m, a, b, \delta > 0$. Here the interpretation is the following. Actions A and B generate both market and non-market outcomes. This could be because action A is intrinsically less productive than action B. Referring, for instance, to our example between religious Muslims and non-religious Muslims, it could be that certain religious practices (Ramadan fasting, praying, etc.) may be associated to productive inefficiencies of religious people as compared to non religious ones and that markets internalize partly such effects. For instance, Campante and Yanagizawa-Drott (2015) show that Ramadan fasting may have a significant negative effect on output growth in Muslim countries. Also, the Dinar 2011 Survey[3] indicates a loss on average of two hours work per day during Ramadan. Alternatively the market outcome associated with action A can also be lower than the one associated to action B because of *social discrimination* related to individuals practicing action A (think about actions such as specific clothes like wearing a veil or a burqa) although such action would not induce per se any direct productivity loss. In such a case, δ will be a measure of the extent of social discrimination or penalization associated to individuals choosing action A. At the same time, depending on the cultural trait, actions A and B provide also different non-market outcomes. Specifically, an individual of type a taking action A receives a strictly

[1]PSE and Ecole des Ponts ParisTech, PUC-Rio and CEPR. Email: thierry.verdier@ens.fr.

[2]Monash University, IFN and CEPR. Email: yves.zenou@monash.edu.

[3]see:
http://www.dinarstandard.com/wp-content/uploads/2013/05/2011-Productivity-in-Ramadan-Report.pdf.

positive non-market benefit $a + v^a G^a$ while action B does not generate any positive non-market benefit. Conversely, an individual of type b taking action B receives a strictly positive non-market benefit $b+v^b G^b$ while action A does not give her any positive non-market benefit. Obviously when $a - \delta u^m$ is positive, this payoff structure is equivalent to the one in (5) with $u^a = a - \delta u^m > 0$ and $u^b = \delta u^m + b > 0$ and setting the utility normalization $U^a(B, G^a) = u^m$ and $U^b(A, G^b) = (1-\delta)\, u^m$ instead of zero. Note that under either payoff structure (5) or (A.1), with $a - \delta u^m > 0$, clearly action $x = A$ (resp. $x = B$) is the preferred action for individuals of type $i = a$ (resp. $i = b$). It should be clear that all the results obtained in this paper will be the same under payoffs (5) or (A.1).

In the main text, we will use payoffs (5) but we can interpret some policy results using payoffs (A.1) since they can give some insights on the productivity and discrimination of ethnic minorities. Indeed, if we consider Propositions 1 and 2, with payoffs (A.1), an alternative "integration" policy could also be to increase the opportunity cost δu^m of the action linked to trait a. This could be done, for instance, by increasing the market value u^m that the minority individuals can get when adopting the behavior corresponding to the "mainstream" culture. Although sometimes difficult to implement, repressive policies against social or economic discrimination or affirmative action policies promoting visible minorities may help achieve this outcome. Note, as well, that an increase in δ, the social discount or penalty factor associated to the action linked to trait a will also negatively affect the family incentives to transmit trait a in the minority population.

References

[1] Campante, F. and D. Yanagizawa-Drott (2015), "Does religion affect economic growth and happiness? Evidence from Ramadan," *Quarterly Journal of Economics* 130, 615-658.

B Online Appendix B: Extensions of the benchmark case

B.1 Online Appendix B.1: Non-rivalry of the community-specific public good

Consider now that the public good $G^a = G$ of community a has some *non-rivalry* dimension so that individuals of type b also derive some benefits from to the provision of G. An example of this is the organization of religious and cultural events that have positive recreative spillovers (fireworks, public entertainments, music shows, street marching, neighborhood parties) on individuals who do not necessarily actively participate, or personally share the cultural views of the cultural community that promotes specifically such events. A simple way to capture this feature is to modify the benchmark payoff specification (5) (with $G^b = 0$ as it is assumed in the main text) in the following way:

$$\begin{aligned}
U^a(A, G) &= u^a + (v^a + \lambda_A)G \\
U^a(B, G) &= \lambda_B G \\
U^b(B, G) &= u^b + \lambda_B G \\
U^b(A, G) &= \lambda_A G
\end{aligned} \tag{B.1}$$

where $\lambda_A, \lambda_B \geq 0$ capture the extent of the spillover effect of the public good G when action A or action B is exerted. As before, individuals of type a obtain some specific benefit $v^a G$ associated with the public good, when they choose the preferred action (i.e. action A) associated to trait a. On top of this, there is also a spillover effect $\lambda_x G$ that is enjoyed by all individuals and which, eventually, depends on the nature of the action $x \in \{A, B\}$ that is chosen.

It is natural to assume $\Delta\lambda = \lambda_A - \lambda_B \geq 0$, that is the strength of the spillover of G is stronger for action A (the action that an individual of type a intrinsically prefers) than for action B (the action that an individual of type b intrinsically prefers). In other words, while the provision of the public good G produces some spillover effects beyond the strict boundaries of the community of individuals of type a, these spillovers are still more effective when one undertakes the behavior preferred by such individuals. Moreover, we assume that $\Delta\lambda\overline{G} < u^b$, so that an individual with cultural trait b still always prefers to undertake action B even if this action is less "complementary" to the public good than action A.

It is then easy to see that the incentives to transmit the different cultural traits inside the families can be written as

$$\Delta V^a = U^a(A, G) - U^a(B, G) = (v^a + \Delta\lambda)\, G + u^a$$

and

$$\Delta V^b = U^b(B,G) - U^b(A,G) = u^b - \Delta\lambda G$$

Proceeding as for the benchmark model, the transmission of preferences can then be described by the following equation:

$$\overset{\bullet}{q}_t = q_t(1-q_t)\left[(1-q_t)\left[u^a + (v^a + \Delta\lambda)\,G_t\right] - q_t\left(u^b - \Delta\lambda G_t\right)\right]$$

or

$$\overset{\bullet}{q}_t = q_t(1-q_t)\left[(1-q_t)\left(u^a + v^a G_t\right) - q_t u^b + \Delta\lambda G_t\right] \tag{B.2}$$

As in the benchmark model, we start in a long-run cultural situation when the leader is never active ($G = 0$), i.e. $q_0 = q(0)$, where the initial fraction of individual with trait a is equal to the steady-state without a cultural leader, i.e. $q_0 = u^a/(u^a + u^b)$. Because of $\Delta\lambda\overline{G} < u^b$, the cultural steady state $q^*\left(\overline{G}\right)$ associated to the dynamics (B.2) with the full leader's public good capacity $G_t = \overline{G}$, is always interior.[4]

Assuming again, for simplicity, that $\beta = v^a = 1$. Then, the program of the cultural leader of group a is given by:

$$\begin{aligned} &\max_{0\le G_t\le \overline{G}} \int_0^\infty e^{-\rho t}\left(W^a q_t - cG_t\right)dt \\ s.t.\quad \overset{\bullet}{q}_t &= q_t(1-q_t)\left[(1-q_t)\left(u^a + G_t\right) - q_t u^b + \Delta\lambda G_t\right] \\ &s.t.\quad q_0 = q(0) \text{ is given} \end{aligned} \tag{B.3}$$

We proceed as for the benchmark model. After substitution of the control function, and using (B.2), we obtain:

$$\max_{0\le G_t\le \overline{G}} \int_0^\infty e^{-\rho t}\left(W^a q_t - c\left[\frac{\overset{\bullet}{q}_t}{q_t(1-q_t)\left(1+\Delta\lambda - q_t\right)} - \left(\frac{u^a - (u^a+u^b)q_t}{1+\Delta\lambda - q_t}\right)\right]\right)dt$$

Now consider

$$I = \int_0^\infty e^{-\rho t}\left[\frac{\overset{\bullet}{q}_t}{q_t(1-q_t)\left(1+\Delta\lambda - q_t\right)}\right]dt$$

and denote the function $V(q)$ as:

$$V(q) = \int_{q_0}^{q} \frac{du}{u(1-u)\left(1+\Delta\lambda - u\right)}$$

[4]Indeed

$$q^*\left(\overline{G}\right) = \frac{u^a + \Delta\lambda\overline{G}}{u^a + \overline{G} + u^b},$$

which is always less than 1.

We immediately have that

$$\frac{1}{q_t(1-q_t)\,(1+\Delta\lambda-q_t)} = \frac{1}{(1+\Delta\lambda)}\frac{1}{q_t} + \frac{1}{\Delta\lambda}\frac{1}{(1-q_t)} - \frac{1}{\Delta\lambda(1+\Delta\lambda)}\frac{1}{(1+\Delta\lambda-q_t)}$$

By integration, we obtain:

$$V(q) = Cte + \frac{1}{(1+\Delta\lambda)}\ln q - \frac{1}{\Delta\lambda}\ln(1-q) + \frac{1}{\Delta\lambda(1+\Delta\lambda)}\ln(1+\Delta\lambda-q_t)$$

with $V(q_0) = 0$. Integration by parts then leads to:

$$\begin{aligned} I &= \int_0^\infty e^{-\rho t}\frac{\dot{q}_t}{q_t(1-q_t)\,(1+\Delta\lambda-q_t)} \\ &= \left[e^{-\rho t}V(q_t)\right]_0^\infty + \rho\int_0^\infty e^{-\rho t}V(q_t)dt \\ &= \rho\int_0^\infty e^{-\rho t}V(q_t)dt \end{aligned}$$

The optimal control problem (B.3) of the leader is therefore equivalent to

$$\max_{0\le G_t\le \overline{G}}\int_0^\infty e^{-\rho t}R(q_t)dt$$

where

$$R(q_t) = W^a q_t + c\left[\frac{u^a-(u^a+u^b)q_t}{1+\Delta\lambda-q_t}\right] - c\rho V(q_t)$$

Now observe that

$$R'(q) = W^a - c\frac{(u^a+u^b)\,(1+\Delta\lambda)-u^a}{(1+\Delta\lambda-q_t)^2} - \frac{c\rho}{q_t(1-q_t)\,(1+\Delta\lambda-q_t)}$$

Consider the function

$$\Phi(q) = \frac{H}{(1+\Delta\lambda-q_t)^2} + \frac{\rho}{q_t(1-q_t)\,(1+\Delta\lambda-q_t)}$$

with

$$H = (u^a+u^b)\,(1+\Delta\lambda) - u^a > 0$$

Inspection of $\Phi(q)$ shows that it has the shape shown in Figures B1(a) and B1(b).

[Insert Figures B1(a) and B1(b) here]

Formally, it satisfies the following properties:

(i) First, $\lim_{q\to 0}\Phi(q) = +\infty$ and $\lim_{q\to 1}\Phi(q) = +\infty$.

(ii) Second, there is a unique value $\widehat{q} \in (0,1)$ such that $\Phi(q)$ is decreasing for $q \in (0, \widehat{q})$ and increasing for $q \in (\widehat{q}, 1)$, reaching therefore its minimum at $\Phi(\widehat{q})$.

From (i) and (ii), it follows that, for $W^a < \Phi(\widehat{q})$, then $R'(q) < 0$ and $R(q)$ is decreasing for $q \in (0,1)$. In particular $R(q)$ is continuous and decreasing for $q \in \left[q_0, q^*(\overline{G})\right]$. On the other hand, for $W^a > \Phi(\widehat{q})$, then there exist q_L and q_H with $0 < q_L < \widehat{q} < q_H < 1$, such that $R'(q) < 0$ for $q \in (0, q_L) \cup (q_H, 1)$ and $R'(q) > 0$ for $q \in (q_L, q_H)$. Therefore, $R(q)$ has a local minimum at q_L and a local maximum at q_H .

The shape of the function $R(q)$ is displayed in Figure B2 and it can be seen that it has the same shape as in the benchmark model (see Figure 1). Thus, one can apply the MRAP method of Spence and Starrett (1975) in exactly the same way. The results are qualitatively similar to those described in Proposition 1 where now $\overline{W} = \Phi(\widehat{q})$. In other words, the optimal decision of the amount of public good G^a provided by the leader is described by Proposition 1. Note, in particular, that when $\Delta\lambda = 0$ (i.e. $\lambda_A = \lambda_B$ so that the strength of the spillovers is the same for both communities), we obtain exactly the results described in Proposition 1. In that case, the externality of the public good is not associated to one particular action. Therefore, the incentives to transmit one trait do not depend on it.

It is also immediate to see that an increase in $\Delta\lambda$, the complementarity between the public good and the optimal behavior (i.e. action A) of cultural trait a, shifts up the marginal benefit $R'(q)$ and shifts down the minimum marginal cost $c\Phi(\widehat{q})$. This makes it more likely for the cultural leader to be active and supply the public good G^a. Indeed, it will be profitable for the leader to supply G at a lower minimum threshold $\overline{W}$ of cultural rents, and at a lower minimum size $\widehat{q}^*$ of the initial fraction of the population q_0.

In words, an increase in the difference between the strength of spillovers between the two communities shifts up the marginal benefit for the cultural leader of providing the public good G^a and shifts down the minimum marginal cost. This implies that the cultural leader is more likely to be active when the spillover benefits of the public good G^a are more specific to individuals from his own group. This also suggests that, if the leader can select to some extent the characteristics of the public good he wants to provide, he will choose a good that generates more "community-specific" spillover effects, which implies that the results will be closer to that of our benchmark model.

[*Insert Figure B2 here*]

B.2 Online Appendix B.2: The leader maximizes socialization rents

We may consider that the objective function of the leader interacts with the socialization efforts of members of her community. More specifically, consider the case where the leader's preferences are such that she maximizes the following rents (using (14) and assuming for simplicity that $\beta = v^a = 1$):

$$q_t d_t^a = q_t(1-q_t)(u^a + G^a)$$

These rents are just the fraction of the population of type a times the optimal effort of type$-a$ parents. In other words, the objective function of the leader is now given by: $\int_0^\infty e^{-\rho t}\left(q_t d_t^a - cG_t^a\right)dt$. Note first that, after substitution of the control function, using (9), we obtain:

$$\begin{aligned} q(1-q)(u^a+G^a) &= q(1-q)u^a + q(1-q)\left[\frac{\dot{q}}{q(1-q)^2} + \frac{qu^b}{(1-q)} - u^a\right] \\ &= q^2u^b + \frac{\dot{q}}{(1-q)} \end{aligned}$$

The objective function $q_t d_t^a - cG_t^a$ of the leader can now be written as:

$$\begin{aligned} q(1-q)(u^a+G^a) - cG^a &= q^2u^b + \frac{\dot{q}}{(1-q)} - c\left[\frac{\dot{q}}{q(1-q)^2} + \frac{qu^b}{(1-q)} - u^a\right] \\ &= q^2u^b - \frac{cu^bq}{(1-q)} + cu^a - c\frac{\dot{q}}{q(1-q)^2} + \frac{\dot{q}}{(1-q)} \end{aligned}$$

Therefore, the leader maximizes the following function:

$$\int_0^\infty e^{-\rho t}\left[q^2u^b - \frac{cu^bq}{(1-q)} + cu^a - c\frac{\dot{q}}{q(1-q)^2} + \frac{\dot{q}}{(1-q)}\right]dt \tag{B.4}$$

Noting that

$$\frac{1}{q(1-q)^2} = \frac{1}{q} + \frac{1}{1-q} + \frac{1}{(1-q)^2}$$

and integration by parts of the term with $\dot{q}_t$ in (B.4) provides that the optimal control problem of the cultural leader is then equivalent to the following program (up to a constant):

$$\max_{0 \le G_t^a \le \overline{G}^a} \int_0^\infty e^{-\rho t} R(q_t) dt$$

where

$$R(q_t) = q_t^2 u^b - \rho \log(1-q_t) - c\left(\frac{u^b q_t + \rho}{1-q_t}\right) - c\rho \log\left(\frac{q_t}{1-q_t}\right)$$

It is easily verified that:

$$R'(q_t) = 2q_t u^b + \frac{\rho}{1-q_t} - c\left[\frac{\rho + u^b q_t}{q_t(1-q_t)^2}\right]$$

This can be rewritten as

$$q_t(1-q_t)^2 R'(q_t) = 2q_t^2(1-q_t)^2 u^b + \rho q_t(1-q_t) - c\left[\rho + u^b q_t\right]$$

Note that $q_t(1-q_t)^2 R'(q_t)$ is a function that is the difference between two terms $\Phi(q_t) = 2q_t^2(1-q_t)^2 u^b + \rho q_t(1-q_t)$ and $\Gamma(q_t) = c\left[\rho + u^b q_t\right]$. It is easy to see that the function $\Phi(q_t)$ is increasing in q_t if and only if $q_t \in [0, 1/2]$, reaching therefore its maximum at $\Phi(1/2) = \frac{1}{4}(\frac{u^b}{2} + \rho)$ and such that $\Phi(0) = \Phi(1) = 0$. While $\Gamma(q_t)$ is linear increasing in q_t with $\Gamma(1/2) = c\left[\rho + \frac{u^b}{2}\right]$. From this it follows easily that:

(i) for $c > \frac{1}{4}$, one has $R'(q_t) = \Phi(q_t) - \Gamma(q_t) < 0$ for all $q_t \in [0, 1]$ while

(ii) for $c \leq 1/4$ there are two values q_L and q_H with $q_L \leq 1/2 \leq q_H$ and such that $R'(q_t) \leq 0$ for $q_t \in [0, q_L] \cup [q_H, 1]$ and $R'(q_t) \geq 0$ for $q_t \in [q_L, q_H]$.

Hence for $c \geq \frac{1}{4}$ $R(q_t)$ is always decreasing in $q_t \in [0, 1]$ while for $c < 1/4$, $R(q_t)$ has a local minimum at q_L and a local maximum at q_H. The function $R(q_t)$ takes therefore the same shape as in Figure 1, and the analysis of the optimal degree of cultural participation of the leader is very similar to the benchmark case where the leader is maximizing a rent proportional to the size of her group.

B.3 Online Appendix B.3: The leader of type a maximizes the utility of the population of type a

We may consider that the objective function of the leader is to maximize the expected utility of individuals of the same type (i.e. of type a). In that case, instead of (12), the leader will maximize the following function (normalizing $v^a = 1$ and $\beta = 1$):

$$\max_{0 \leq G_t^a \leq \overline{G}^a} \int_0^\infty e^{-\rho t}\left[q_t\left(u^a + G_t^a\right) - cG_t^a\right] dt$$

$$s.t. \quad \overset{\bullet}{q}_t = q_t(1-q_t)\left[(1-q_t)\left(u^a + G_t^a\right) - q_t u^b\right]$$

$$s.t. \quad q_0 = q(0) \text{ given}$$

Note first that, using (9), we obtain:

$$G_t^a = \frac{\overset{\bullet}{q}_t}{q_t(1-q_t)^2} + \frac{q_t}{(1-q_t)} u^b - u^a$$

By substituting G_t^a in the control function, the objective function is given by:

$$\begin{aligned} R(q_t) &= \frac{\overset{\bullet}{q}_t}{(1-q_t)^2} - c\frac{\overset{\bullet}{q}_t}{q_t(1-q_t)^2} + \frac{q_t^2}{(1-q_t)}u^b - c\frac{q_t}{(1-q_t)}u^b + cu^a \\ &= cu^a + \frac{(q_t-c)}{(1-q_t)}q_t u^b + \frac{(q_t-c)}{q_t(1-q_t)^2}\overset{\bullet}{q}_t \end{aligned}$$

Therefore, the leader maximizes the following function:

$$\int_0^\infty e^{-\rho t}\left[cu^a + \frac{(q_t-c)}{(1-q_t)}q_t u^b + \frac{(q_t-c)}{q_t(1-q_t)^2}\overset{\bullet}{q}_t\right]dt \tag{B.5}$$

It is easily verified that

$$\frac{(q_t-c)}{q_t(1-q_t)^2} = -\frac{c}{q_t} - \frac{c}{1-q_t} + \frac{1-c}{(1-q_t)^2}$$

Then integration by parts of the term with $\overset{\bullet}{q}_t$ in (B.5) provides easily that, up to some constant, our optimal control problem collapses to

$$\max_{0\leq G_t^a\leq \overline{G}^a}\int_0^\infty e^{-\rho t}R(q_t)dt$$

where

$$R(q_t) = c\,u^a + \frac{q_t^2 u^b+\rho}{(1-q_t)} - \frac{c\left(q_t u^b+\rho\right)}{(1-q_t)} - c\rho\log\left(\frac{q_t}{1-q_t}\right)$$

Let us study $R(q_t)$. We have:

$$R'(q_t) = \frac{q_t u^b(2-q_t)+\rho}{(1-q_t)^2} - \frac{c\left(\rho+q_t u^b\right)}{q_t(1-q_t)^2}$$

and thus

$$(1-q_t)^2 R'(q_t) = \left[q_t u^b(2-q_t)+\rho\right] - \frac{c\left(\rho+q_t u^b\right)}{q_t}$$

Note that $(1-q_t)^2 R'(q_t)$ is a function that is the difference between two terms $\Omega(q_t) = q_t u^b(2-q_t)+\rho$ and $\Theta(q_t) = \frac{c\left(\rho+q_t u^b\right)}{q_t}$. It is easy to see that the function $\Omega(q_t)$ is increasing for all $q_t \in [0,1]$, and such that $\Phi(0) = \rho$, $\Phi(1) = \rho + u^b$. Moreover, $\Theta(q_t)$ is a decreasing in q_t with $\Theta(0) = +\infty$ and $\Theta(1) = c\left[\rho+u^b\right]$. From this it follows easily that, given that $c < 1$, there exists a unique value $q_M \in]0,1[$ such that $R'(q_t) \leq 0$ if and only if $q_t \leq q_M$. Hence $R(q_t)$ is having a minimum at q_M with $R(0) = R(1) = +\infty$. For $q_t \in \left[q(0), q(\overline{G}^a)\right]$, the function $R(q_t)$ takes the shape as in Figure 1, and it is easy to see that the analysis of the optimal degree of cultural participation of the leader is very similar to the benchmark case where the leader is maximizing a rent proportional to the size of her group.

C Online Appendix C: Club Good with community members' contributions

In the benchmark model, we assumed that the cultural leader was financing himself the cost of the provision of the group-specific public good $G = G^a$. This could be justified if, for example, the leader is a foreign power that has a clear objective of building up a group of individuals favorable to its specific religious or political positions. An interesting example of this is the case in Kosovo of clerics funded by money from Saudi Arabia to promote the emergence of radical Islamic communities promoting active members for ISIS (New York Times article 'How Kosovo was Turned into Fertile Ground for ISIS' May 22, 2016). While in the case of a foreign power, the contribution and possibly the choice of public goods comes from outside the group, in many cases the financing of the public good comes from contributions collected on members of the community. In this Appendix, we extend the benchmark model to capture these features.

Consider now that the public good G is financed by a contribution $\tau^a \in [0, \overline{\tau}]$ on community members. The utility functions of the different individuals are now by the following payoffs:

$$\begin{aligned}
U^a(A, G) &= v^a G - \tau^a \\
U^a(B, G) &= 0 \\
U^b(B, G) &= 0 \\
U^b(A, G) &= 0 - \tau^a
\end{aligned} \tag{C.1}$$

where, for simplicity, $u^a = u^b = 0$. Note that taking action A corresponds to participating in the club activity of group a and therefore to pay the contribution τ^a associated to this participation. Clearly individuals of type a have a valuation $v^a G$ of the club activity, which is increasing in G while individuals of type b do not value this club good. As a consequence, the optimal action x^a of individuals of type a will be action $x^a = A$ when $v^a G - \tau^a \geq 0$, and action $x^a = B$ otherwise. The optimal action x^b of individuals of type b will always be action $x^b = B$ (non participation to the club good). The incentives to transmit the different cultural traits inside the families can then be written as

$$\Delta V^a = U^a(x^a, G) - U^a(x^b, G) = \max\{v^a G - \tau^a, 0\}$$

and

$$\Delta V^b = U^b(x^b, G) - U^b(x^a, G) = \mathbf{1}\left(v^a G - \tau^a\right) \tau^a$$

where $\mathbf{1}(x)$ is the indicator function such that $\mathbf{1}(x) = 1$ when $x \geq 0$ and $\mathbf{1}(x) = 0$ otherwise.

The budget constraint of the good G at time t can be written as

$$cG_t = \tau_t^a q_t \tag{C.2}$$

Thus

$$\Delta V^a = \tau_t^a \max\left\{\frac{v^a}{c} q_t - 1, 0\right\} \text{ and } \Delta V^b = \mathbf{1}\left(\frac{v^a}{c} q_t - 1\right)\tau_t^a \tag{C.3}$$

To avoid non-trivial dynamics, assume that $\theta = v^a/c > 1$ (otherwise participation in the club good will never be optimal). Proceeding as in the benchmark model, the transmission of preferences can be described by the following equation:

$$\overset{\bullet}{q}_t = q_t(1-q_t)\tau_t^a\left[(1-q_t)\max\{\theta q_t - 1, 0\} - q_t \mathbf{1}(\theta q_t - 1)\right]$$

We consider the case when $q_0 > 1/\theta$ so that the size of the community is large enough to induce some club participation initially at least from a static point of view.

First, note that, whenever $q_0 > 1/\theta$, for all contribution $\tau_t^a \geq 0$, the induced cultural dynamics of the community is such that $q_t \geq 1/\theta$. Indeed, the cultural dynamics is given by:

$$\overset{\bullet}{q}_t = \begin{cases} 0 & \text{when } q_t < 1/\theta \\ q_t(1-q_t)\tau_t^a\left[\theta(1-q_t)q_t - 1\right] & \text{when } q_t \geq 1/\theta \end{cases} \tag{C.4}$$

The steady states for $q_0 > 1/\theta$ do not depend on τ_t^a, which only affects the speed of convergence. When $\theta < 4$, we always have $\overset{\bullet}{q}_t < 0$ for all $q_t \in [1/\theta, 1)$ as long as $\tau_t^a > 0$. Therefore, if $\tau_t^a > 0$, for all t, then q_t converges towards $1/\theta$ from above (the steady state $q = 1$ is unstable). Otherwise, when $\tau_{t_0}^a = 0$ at some time t_0, then the dynamics stops at $q_{t_0} > 1/\theta$.

When $\theta > 4$, there are three steady states for the cultural dynamics: $q_{\min}$, $q_{\max}$ and 1 such that

$$1/\theta < q_{\min} < q_{\max} < 1$$

and only $q_{\max}$ is stable. In such a case, if $\tau_t^a > 0$ at all time, then the system converges to $1/\theta$ when $q_0 \in [1/\theta, q_{\min}]$ and to $q_{\max}$ when $q_0 > q_{\min}$. Whenever $\tau_t^a = 0$, the system stops at some value $q_{t_0} > 1/\theta$. In the sequel, we consider that $\theta > 4$ so that the utility of the club good is high enough relative to its costs, which implies that a positive active community can eventually be sustained in the long run at some value $q_{\max} > 1/\theta$.

We have the following result:

Proposition C1: *Assume that* $\theta > 4$ and $q_0 > 1/\theta$, where $\theta = v^a/c > 1$.

(i) When $q_0 \in (1/\theta, \widetilde{q})$, the cultural leader does not produce the club good G and the size q_t of the cultural group a remains stationary at $q_t = q_0$ for all t.

(ii) When $q_0 \in (\widetilde{q}, q_{\min})$, then the cultural leader produces the club good G with the maximum contribution rate $\overline{\tau}$ and stops the production of the club good at some t_0 such that $q_{t_0} = \widetilde{q}$. The size q_t of the cultural group a decreases from q_0 to $\widetilde{q}$ and then stays stationary at $q_t = \widetilde{q}$ for all $t \geq t_0$.

(iii) When $q_0 \in (q_{\min}, q_{\max})$, then the cultural leader produces the club good G at all t with the maximum contribution rate $\overline{\tau}$. The size q_t of the cultural group a grows and converges towards the steady-state $\lim_{t\to\infty} q_t = q_{\max}$.

Proof of Proposition C1: The cultural leader maximizes a discounted sum of an average of some cultural rents proportional to the size of the community of type a, i.e. $W^a q_t$, and the sum of the utility of the members of the community

$$\begin{aligned} W(q_t, \tau_t^a) &= \mu W^a q_t + (1-\mu) q_t \max\{v^a G^a - \tau_t^a, 0\} \\ &= \mu W^a q_t + (1-\mu) q_t \tau_t^a \max\{\theta q_t - 1, 0\} \end{aligned}$$

The intertemporal payoff of the leader is given by:

$$\int_0^\infty e^{-\rho t} W(q_t, \tau_t^a) dt$$

Obviously

$$W(q_t, \tau_t^a) = \begin{cases} \mu W^a q_t \text{ and } \dot{q}_t = 0 & \text{when} \quad q_t < 1/\theta \\ \mu W^a q_t + (1-\mu) \frac{\dot{q}_t}{(1-q_t)} \frac{(\theta q_t - 1)}{[\theta(1-q_t)q_t - 1]} & \text{when} \quad q_t \geq 1/\theta \text{ and } \tau_t^a > 0 \end{cases}$$

Define the functions:

$$\begin{aligned} H(q) &= \frac{1}{1-q} \frac{(\theta q - 1)}{[\theta(1-q)q - 1]} \quad \text{for } q \geq 1/\theta \\ V(q) &= \int_{1/\theta}^{q} H(u) du \end{aligned}$$

It is easily verified that

$$H(q) = \frac{A}{1-q} + \frac{B}{q_{\max} - q} + \frac{C}{q - q_{\min}}$$

where

$$A = (1-\theta)\theta, \; B = \frac{\theta\left[(\theta - 1) q_{\max} - 1\right]}{q_{\max} - q_{\min}}, \; C = \frac{\theta\left[(\theta - 1) q_{\min} - 1\right]}{q_{\max} - q_{\min}}$$

and $A < 0$, $B > 0$ and $C > 0$ (as $\frac{1}{\theta - 1} < q_{\min} < q_{\max}$ for $\theta > 4$). Hence

$$V(q) = \begin{cases} Cte - A\ln(1-q) - B\ln(q_{\max} - q) + C\ln(q_{\min} - q)) & \text{for} \quad q \in (1/\theta, q_{\min}) \\ Cte - A\ln(1-q) - B\ln(q_{\max} - q) + C\ln(q - q_{\min}) & \text{for} \quad q \in (q_{\min}, q_{\max}) \end{cases}$$

This implies that

$$\lim_{q \to q_{\min}^-} V(q) = -\infty, \lim_{q \to q_{\min}^+} V(q) = -\infty \text{ and } \lim_{q \to q_{\max}-} V(q) = +\infty$$

Moreover, because of the sign of $H(q)$ on the interval $[1/\theta, q_{\max}]$, $V(q)$ is decreasing on the interval $q \in (1/\theta, q_{\min})$ and increasing on the interval $q \in (q_{\min}, q_{\max})$. Moreover there is a unique $q^r \in (q_{\min}, q_{\max})$ such that $V(q^r) = 0$.

Consider now the integral:

$$I = (1-\mu) \int_0^\infty e^{-\rho t} \left[\frac{\dot{q}_t}{(1-q_t)} \frac{(\theta q_t - 1)}{[\theta(1-q_t)q_t - 1]} \right] dt$$

Integration by parts gives:

$$\begin{aligned} I &= (1-\mu) \int_0^\infty e^{-\rho t} \dot{q}_t H(q_t) dt \\ &= (1-\mu) \left[e^{-\rho t} V(q_t) \right]_0^\infty + (1-\mu) \int_0^\infty \rho e^{-\rho t} V(q_t) dt \\ &= (1-\mu) \left[\lim_{t \to \infty} e^{-\rho t} V(q_t) - V(q_0) \right] + (1-\mu) \int_0^\infty \rho e^{-\rho t} V(q_t) dt \end{aligned}$$

Note that, for $q_0 < q_{\min}$, $q_t \in (1/\theta, q_0)$ and $V(q_t)$ is absolutely bounded and $\lim_{t \to \infty} e^{-\rho t} V(q_t) = 0$. Now, for $q_0 > q_{\min}$, two cases can occur:

Either (i) there exists a t such that $\tau^a = 0$ and the cultural dynamics stop. Then again $\lim_{t \to \infty} q_t < q_{\max}$. Therefore $V(q_t)$ is again absolutely bounded and $\lim_{t \to \infty} e^{-\rho t} V(q_t) = 0$.

Or (ii) $\tau_t^a > 0$ for all t. Then, $\lim_{t \to \infty} q_t = q_{\max}$, and

$$\lim_{t \to \infty} V(q_t) \sim \lim_{t \to \infty} -B \log(q_{\max} - q_t) = +\infty$$

However, we have $\lim_{t \to \infty} e^{-\rho t} V(q_t) = 0$. Indeed denote $G(t) = -e^{-\rho t} \log(q_{\max} - q_t)$. Then,

$$\begin{aligned} \dot{G}(t) &= -\rho G(t) + \frac{\dot{q}_t e^{-\rho t}}{q_{\max} - q_t} \\ \dot{G}(t) + \rho G(t) &= e^{-\rho t} q_t (1-q_t) \tau_t^a (q_t - q_{\min}) \end{aligned}$$

Thus

$$\int_0^T \left[\dot{G}(u) + \rho G(u) \right] e^{\rho u} du < \int_0^T q_t (1-q_t)(q_t - q_{\min}) dt < \frac{1}{4} (q_{\max} - q_{\min})$$

As a result,

$$G(T) e^{\rho T} < G(0) + \frac{1}{4} (q_{\max} - q_{\min})$$

and $\lim_{T\to\infty} G(T) = 0$. From this, we also conclude that

$$\lim_{t\to\infty} e^{-\rho t} V(q_t) \sim B \lim_{t\to\infty} G(t) = 0$$

Thus, for all initial conditions $q_0 > 1/\theta$, the intertemporal payoff of the cultural leader can be rewritten as

$$\begin{aligned}\int_0^\infty e^{-\rho t}\left[\mu W^a q_t + (1-\mu) q_t \tau_t^a (\theta q_t - 1)\right] dt &= \int_0^\infty e^{-\rho t} \mu W^a q_t dt + I \\ = \int_0^\infty e^{-\rho t}\left[\mu W^a q_t + (1-\mu)\rho V(q_t)\right] dt &\end{aligned}$$

Thus the problem of the cultural leader collapses to

$$\max_{q_t} \int_0^\infty e^{-\rho t} R(q_t) dt \tag{C.5}$$

where

$$R(q) = \mu W^a q + (1-\mu)\,\rho V(q)$$

This function is such that

$$R'(q) = \mu W^a + (1-\mu)\,\rho H(q)$$

Denote

$$\Delta\,(q) = \mu W^a (1-q)\left[\theta q(1-q) - 1\right] + (1-\mu)\,\rho(\theta q - 1)$$

Then

$$R'(q) = \frac{-\Delta\,(q)}{(1-q)\left[1 - q(1-q)\theta\right]}$$

As we know, when $\theta > 4$, the function $\mu W^a \Theta\,(q) = (1-q)\left[1 - q(1-q)\theta\right]$ has three zeros: $q_{\min}$, $q_{\max}$ and 1 such that $1/\theta < q_{\min} < q_{\max} < 1$. Moreover, $\Theta\,(q)$ is positive for $q \in [1/\theta, q_{\min}] \cup [q_{\max}, 1]$ (as $\Theta\,(1/\theta) > 0$ and $\Theta'\,(1) < 0$) and negative for $q \in (q_{\min}, q_{\max})$. Also, it is straightforward to see that $\Theta'\,(q) < 0$ for $q \in [1/\theta, q_{\min}]$.

Given that for all $\tau^a \geq 0$ and $q_0 \in [1/\theta, q_{\max}]$, the cultural dynamics q_t remains into the interval $[1/\theta, q_{\max}]$, we can restrict ourselves to consider the optimal trajectories in this support.

It is then a simple matter to see that there exists a unique value $\widetilde{q} \in (1/\theta, q_{\min})$ such that

$$\Delta\,(q) \lesseqgtr 0 \text{ if } q \lesseqgtr \widetilde{q}$$

Indeed, at $q = 1/\theta$, one has $\Delta\,(1/\theta) = -\mu W^a \Theta\,(1/\theta) < 0$ and $\Delta\,(q) = (1-\mu)\,\rho(\theta q - 1) - \Theta\,(q) > 0$ for all $q \in (q_{\min}, q_{\max})$. Moreover, in the interval $(1/\theta, q_{\min})$,

$$\Delta'\,(q) = (1-\mu)\,\rho\theta - \Theta'\,(q) > 0$$

Therefore there exists a unique $\widetilde{q} \in (1/\theta, q_{\min})$ such that $\Delta(q) \lesseqgtr 0$ if $q \lesseqgtr \widetilde{q}$.

From this, we conclude that $R'(q)$ is positive for $q \in [1/\theta, \widetilde{q}]$, negative for $q \in [\widetilde{q}; q_{\min}]$ and $R'(q)$ is again positive for $q \in (q_{\min}, q_{\max})$. This implies that $R(q)$ has one local optimum $\widetilde{q}$ in the interval$[1/\theta, q_{\min}(\theta)]$ and is increasing from $-\infty$ to $+\infty$ in the interval $(q_{\min}, q_{\max})$. The shape of the function $R(q)$ is displayed in Figure C1.

[*Insert Figure C*1 *here*]

We have two cases:

(a) $q_0 \in (1/\theta, q_{\min})$. Then, we know that, for all t, $q_t \in [1/\theta, q_0]$. Indeed as long as $\tau_t^a > 0$ (the club good is active), then q_t is decreasing in t and therefore $q_t \in [1/\theta, q_0]$ and if there exists some time $[t_0, t_1]$ such that $\tau_t^a = 0$, then $q_t = q_{t_0} < q_0$ for all time $t \in [t_0, t_1]$. Thus again $q_t \in [1/\theta, q_0]$. From this, it follows that the function $R(q_t)$ is continuous and bounded at all $t \in [0, +\infty)$. A simple application of the MRAP provides then results (i) and (ii) stated in Proposition C1.

(b) $q_0 \in (q_{\min}, q_{\max})$, then for all t, $q_t \in [q_0, q_{\max})$, q_t is weakly increasing sequence in t. Moreover, the function $R(q)$ is increasing in $q \in [q_0, q_{\max})$ from $R(q_0)$ towards $+\infty$. Then, the optimal solution to (C.5) is to choose a dynamic path $(q_t)_{t=0,\ldots\infty}$ such that we again go as fast as possible towards $\lim_{t\to\infty} q_t = q_{\max}$ with $\tau_t^a = \overline{\tau}$ at all time. Indeed, take any other dynamic path $(q_t')_{t\geq t_1}$with an associated contribution rate sequence $\left(\tau_t^{'a}\right)_{t\in(0,\infty)}$ satisfying (C.4) and denote the first interval $t \in (t_1, t_2)$ such that $\tau_t^{'a} < \overline{\tau}$. Also denote

$$F(q) = q(1-q)\left[\theta(1-q)q - 1\right]$$

and

$$G(q) = \int_{q_{t_1}}^{q} \frac{dv}{F(v)}$$

$G(q)$ is a well defined function for $q \in (q_0, q_{\max})$, increasing in q as $F(.)$ is positive in the interval $(q_0, q_{\max})$. Given the definition of t_1, for $t \leq t_1$, we have $q_t = q_t'$. Then, integrating (C.4) along the dynamic path $(q_t')_{t\geq t_1}$, gives, for all $t > t_1$,

$$\int_{q_{t_1}}^{q_t'} \frac{dv}{F(v)} = \int_{t_1}^{t} \tau_u^{'a} du < \int_{t_1}^{t} \overline{\tau} du = \int_{q_{t_1}}^{q} \frac{dv}{F(v)}$$

Thus, for all $t > t_1$ $G(q_t') < G(q_t)$ and therefore $q_t' < q_t$. Given that $R(.)$ is increasing in q in the interval $(q_0, q_{\max})$, we also have that $R(q'_t) < R(q_t)$ for $t > t_1$. As a consequence

$$\begin{aligned}\int_0^\infty e^{-\rho t} R(q'_t)dt &= \int_0^{t_1} e^{-\rho t} R(q'_t)dt + \int_{t_1}^\infty e^{-\rho t} R(q'_t)dt \\ &= \int_0^{t_1} e^{-\rho t} R(q_t)dt + \int_{t_1}^\infty e^{-\rho t} R(q'_t)dt \\ &< \int_0^{t_1} e^{-\rho t} R(q_t)dt + \int_{t_1}^\infty e^{-\rho t} R(q_t)dt \\ &= \int_0^\infty e^{-\rho t} R(q_t)dt\end{aligned}$$

and the intertemporal value for the cultural leader of the dynamic path (q'_t) induced by the fee sequence $\left(\tau_t^{\prime a}\right)$ is dominated by the path (q_t) induced by $(\tau_t^a)_t = \overline{\tau}$. Again the cultural leader wants to approach the steady state $q_{\max}$ as fast as possible and we obtain result *(iii)* stated in Proposition C1. ▮

Let us summarize the findings of this section. We demonstrate that the results in terms of dynamics are similar to that of the benchmark case. Specifically, in Proposition C1, we show that there is a minimum threshold value $\widetilde{q}$ of the initial population q_0 of individuals of type a, such that, beyond this threshold, the leader will become active and supply the good G.

Different from the case with external financing though, it could be that the active provision of G might not be sufficient to allow the community of individuals with trait a to increase its size. The reason is that the amount of good G_t provided at each period t now depends on the total contributions collected by the group and, therefore, on the fraction q_t of individuals of type a.

More precisely, the provision of the club good G generates immediate benefits to individuals of type a who consume it. As it triggers a difference of consumption between the two cultural traits, it also induces paternalistic parents of both types (a and b) to increase their socialization effort in order to transmit their own trait to their children. Now, the incentive ΔV^a to transmit trait a is related to the net consumption benefit of the club good $(\theta q_t - 1)\,\overline{\tau}$ for an individual of type a (where $\theta = v^a/c > 1$). ΔV^a therefore depends positively on the fraction q_t of individuals contributing to the club good. On the other hand, the socialization incentive ΔV^b to transmit the other trait depends on the access fee $\overline{\tau}$ to the club good, which is the cost perceived by a paternalistic parent of type b who does not value the club good G and expects his child with preference a to pay $\overline{\tau}$ to access this good. When q_t is low enough (but larger than $1/\theta$), ΔV^a is small compared to ΔV^b. As a consequence, parents of type a socialize less their children than parents of type b and, therefore, the cultural dynamics lead to a reduction of the fraction q_t of individuals with trait a in the population.

In part (i) of Proposition C1, we show that when the initial value $q_0 \in (1/\theta, \widetilde{q})$, the leader does not produce the club good G and the size of the group stays at $q_t = q_0$, for all t. It is easily verified that when $q_t > 1/\theta$, the consumption value of the club good becomes positive for individuals of type a. Therefore, a fully *myopic* leader, who cares about the welfare of his community members, would always produce the club good G as soon as $q_t > 1/\theta$. However, as soon as the leader becomes forward looking, the leader knows that, starting at $q_0 \in (1/\theta, \widetilde{q})$, there are too few individuals of type a to finance the good G and that, eventually, q_t will decrease and end up at $1/\theta$ in steady state. This is going to reduce dynamically its rents $W^a q_t$. Also, the consumption value $(\theta q_t - 1)\,\overline{\tau}$ of the club good to community members is positive but quite small along the transition path from q_0 to $1/\theta$. As a result, contrary to the myopic case, it is optimal for the forward-looking leader to remain inactive and to never produce the good G.

Consider now the case where the group starts with a fraction q_0 slightly above the threshold $\widetilde{q}$ (part (ii) of Proposition C1 with $q_0 \in (\widetilde{q}, q_{\min})$). Although the value of q_0 is still not high enough to trigger some positive cultural dynamics of q_t, the discounted intertemporal consumption value of the club good for individuals of type a is now high enough that it may compensate for the loss of cultural rents that the leader incurs along the transition path from q_0 to $1/\theta$. In that case, the cultural leader will provide the club good G. However, the level of club good that the community can finance is not sufficient to permit the diffusion of trait a in the population. The leader, therefore, remains active and provides the public good only for a finite amount of time. The population of individuals of type a declines and consumes G up to the moment where its per capita provision cost is too high compared to its benefit. At this stage, the cultural leader ceases to supply the public good and the cultural dynamic process stops.

Finally, when the initial fraction q_0 is much larger than the threshold $\widetilde{q}$ (part (iii) of Proposition C1 with $q_0 \in (q_{\min}, q_{\max})$), then the cultural leader provides the public good at the maximum possible rate of individual contribution, and this promotes the diffusion of the cultural trait a to a higher long-run steady state level. This is because the leader anticipates that there are enough individuals of type a to finance the good G and that q_t will increase until it reaches its maximum value.

D Online Appendix D: Direct socialization of leaders

D.1 Direct socialization of the leader

In the benchmark model, we assumed the following cultural transmission process. All children, born without defined cultural traits, are first exposed to their parent's trait (*direct* *vertical socialization*) and, if not directly socialized, are subject to *outside socialization*. The latter is such that the child is matched to a *passive* role model randomly chosen from the society. The *active* role model, i.e. the leader, affects this cultural transmission only *indirectly* through her choice of the public good G^a, which, in turn, impacts on $U^a(A, G^a) = u^a + v^a G^a$, the utility of a type$-a$ individual taking action A. The utility $U^a(A, G^a)$ directly affects d^a, the parent a's effort in transmitting her trait.

In this section, we consider a model where the leader *directly* affects the cultural transmission mechanism. For this we amend our previous framework in the following way. Specifically we now assume that children are first exposed to their parent i's trait (with probability d^i) but, when this fails, the child is subject to *outside socialization* so that, with probability γ, she is directly exposed to the leader (of trait a) while, with probability $1-\gamma$, the child is matched to a *passive* role model randomly chosen in the society (i.e. she adopts trait i with probability q^i).

D.1.1 The model with one leader

In this new transmission mechanism, we can write the transition probabilities, for all $i, j \in \{a, b\}$, as follows:

$$
\begin{aligned}
P^{aa} &= d^a + (1-d^a)\left[\gamma + (1-\gamma)\, q_t\right] \\
P^{ab} &= (1-d^a)(1-\gamma)(1-q_t) \\
P^{bb} &= d^b + (1-d^b)(1-\gamma)(1-q_t) \\
P^{ba} &= (1-d^b)\left[\gamma + (1-\gamma)\, q_t\right]
\end{aligned}
$$

Observe that, in these transition probabilities, we assume that, once a child is exposed to a leader, she automatically adopts trait a. The cultural dynamics of q_t is then given by:

$$q_{t+1} = q_t P^{aa} + (1-q_t) P^{ba} = q_t\left[d^a + (1-d^a)\left[\gamma + (1-\gamma)\, q_t\right]\right] + (1-q_t)(1-d^b)\left[\gamma + (1-\gamma)\, q_t\right]$$

As a result, instead of (6), the continuous-time dynamics version of this equation is now given by:

$$\overset{\bullet}{q}_t = (1-q_t)\left(1-d^b\right)\left[\gamma(1-q_t) + q_t\right] - q_t(1-q_t)(1-d^a)(1-\gamma) \tag{D.1}$$

Let us now study the program of the leader. The program of the leader can now be written as:

$$
\begin{aligned}
&\max_{0 \le \gamma \le 1} \int_0^\infty e^{-\rho t}\left(W^a q_t - c\gamma\right) dt \\
s.t. \quad \overset{\bullet}{q}_t \;&=\; (1-q_t)\left(1-d^b\right)\left[\gamma(1-q_t) + q_t\right] - q_t(1-q_t)(1-d^a)(1-\gamma) \qquad \text{(D.2)}
\end{aligned}
$$

Observe that, compared to (13), the objective function of the leader has changed since she now chooses γ, her influence in the society and not the public good G^a as in the previous section. Solving this program is quite complicated. Thus, we assume that the choice d^i of parent i is exogenous. When $\gamma = 0$, this, however, leads to very simple dynamics converging to the corner solutions $q = 0$ when $d^a < d^b$ and to $q = 1$ when $d^a > d^b$. These situations of pure homogenous populations may prevent the application of the MRAP approach for the maximization problem.[5] Therefore, to avoid such situations and to replicate the idea that with endogenous family socialization rates (and cultural substituability) there is always some long-run cultural heterogeneity (with and without leader intervention), we approximate the constant parent socialization rates by a frequency dependent socialization rate in the following way:

$$
\begin{aligned}
d^a_\epsilon(q) &= \begin{cases} d^a \text{ for } \; q < 1-\epsilon \\ 0 \text{ for } q \geq 1-\epsilon \end{cases} \\
d^b_\epsilon(q) &= \begin{cases} d^b \text{ for } \; q > \epsilon \\ 0 \text{ for } q \leq \epsilon \end{cases}
\end{aligned}
$$

With such socialization rates, it is easily verified that the cultural dynamics without leader intervention converge respectively to the interior steady states ϵ when $d^a < d^b$ and $1-\epsilon$ when $d^a > d^b$. For $\epsilon > 0$, we can then apply our MRAP approach as the associated function $R(q)$ is bounded on the interval $[\epsilon, 1-\epsilon]$. We can then recover the insights of optimal leader intervention in the standard case when $\epsilon \to 0$.

With these modified socialization rates and starting with some initial population frequency $q\,(0) \in\,]\epsilon, 1-\epsilon[$, the cultural dynamic system is again given by (D.1). As in Section 3, we have the following result:

Lemma D1

(i) Assume that $d^a = d^b = d$. Then, up to some constant, the optimal control problem (D.2) is equivalent to:

$$\max_{0 \leq \gamma \leq 1} \int_0^\infty e^{-\rho t} R(q_t) dt$$

where

$$R(q_t) = W^a q_t + \frac{c\rho}{(1-d)} \log\left(1 - q_t\right)$$

(ii) Assume that $d^a \neq d^b$. Then, up to some constant, the optimal control problem (13) is equivalent to:

$$\max_{0 \leq \gamma \leq 1} \int_0^\infty e^{-\rho t} R(q_t) dt$$

[5] The function $R(q)$ may become unbounded as one approaches $q = 0$ or $q = 1$.

where

$$R(q_t) = W^a q_t + \frac{c\left(d^a - d^b\right) q_t}{\left[q_t\left(1-d^a\right) + \left(1-q_t\right)\left(1-d^b\right)\right]} - \frac{c\rho}{\left(1-d^a\right)} \log\left[\frac{1-d^b - q_t\left(d^a - d^b\right)}{1-q_t}\right]$$

Proof of Lemma D1: Observe that, from (D.1), we easily obtain:

$$\gamma = \frac{\overset{\bullet}{q}_t - q_t\left(1-q_t\right)\left(d^a - d^b\right)}{\left(1-q_t\right)\left[q_t\left(1-d^a\right) + \left(1-q_t\right)\left(1-d^b\right)\right]} \tag{D.3}$$

Therefore, we have:

$$\begin{aligned} & W^a q_t - c\gamma \\ = \ & W^a q_t + \frac{c q_t\left(d^a - d^b\right)}{\left[q_t\left(1-d^a\right) + \left(1-q_t\right)\left(1-d^b\right)\right]} - \frac{c\overset{\bullet}{q}_t}{\left(1-q_t\right)\left[q_t\left(1-d^a\right) + \left(1-q_t\right)\left(1-d^b\right)\right]} \end{aligned}$$

As a result, the program (D.2) can be written as:

$$\begin{aligned} & \max_{0\leq\gamma\leq1} \int_0^\infty e^{-\rho t}\left(W^a q_t + c\left(d^a - d^b\right)\frac{q_t}{\left[q_t\left(1-d^a\right) + \left(1-q_t\right)\left(1-d^b\right)\right]}\right) dt \\ & -c \max_{0\leq\gamma\leq1} \int_0^\infty e^{-\rho t}\left(\frac{\overset{\bullet}{q}_t}{\left(1-q_t\right)\left[q_t\left(1-d^a\right) + \left(1-q_t\right)\left(1-d^b\right)\right]}\right) \end{aligned}$$

Observe that

$$\begin{aligned} & \frac{1}{\left(1-q_t\right)\left[q_t\left(1-d^a\right) + \left(1-q_t\right)\left(1-d^b\right)\right]} \\ = \ & \frac{1}{\left(1-q_t\right)\left(1-d^a\right)} - \frac{\left(d^a - d^b\right)}{\left[1-d^b - q_t\left(d^a - d^b\right)\right]\left(1-d^a\right)} \end{aligned}$$

Thus

$$\begin{aligned} & \int_0^\infty e^{-\rho t}\left(\frac{\overset{\bullet}{q}_t}{\left(1-q_t\right)\left[q_t\left(1-d^a\right) + \left(1-q_t\right)\left(1-d^b\right)\right]}\right) \\ = \ & \frac{1}{\left(1-d^a\right)} \int_0^\infty e^{-\rho t}\frac{\overset{\bullet}{q}_t}{\left(1-q_t\right)} dt - \frac{\left(d^a - d^b\right)}{\left(1-d^a\right)} \int_0^\infty e^{-\rho t}\frac{\overset{\bullet}{q}_t}{1-d^b - q_t\left(d^a - d^b\right)} dt \end{aligned}$$

which after integration provides

$$\frac{1}{\left(1-d^a\right)} \log\left[\frac{1-q_0}{1-d^b - q_0\left(d^a - d^b\right)}\right] + \frac{\rho}{\left(1-d^a\right)} \int_0^\infty \log\left[\frac{1-d^b - q_t\left(d^a - d^b\right)}{1-q_t}\right] e^{-\rho t} dt$$

As a result, the objective function of the leader to be maximized can be written as:

$$\int_0^\infty e^{-\rho t} W^a q_t dt + c\left(d^a - d^b\right) \int_0^\infty e^{-\rho t} \frac{q_t}{\left[q_t\left(1-d^a\right)+\left(1-q_t\right)\left(1-d^b\right)\right]} dt$$
$$-\frac{c}{\left(1-d^a\right)} \log\left[\frac{1-q_0}{1-d^b-q_0\left(d^a-d^b\right)}\right] - \frac{c\rho}{\left(1-d^a\right)} \int_0^\infty \log\left[\frac{1-d^b-q_t\left(d^a-d^b\right)}{1-q_t}\right] e^{-\rho t} dt$$

(i) Assume that $d^a = d^b = d$. This objective function simplifies to:

$$\int_0^\infty e^{-\rho t} W^a q_t dt - \frac{c}{(1-d)} \log\left(\frac{1-q_0}{1-d}\right) - \frac{c\rho}{(1-d)} \int_0^\infty \log\left(\frac{1-d}{1-q_t}\right) e^{-\rho t} dt$$
$$= \int_0^\infty e^{-\rho t} \left[W^a q_t + \frac{c\rho}{(1-d)} \log\left(1-q_t\right)\right] dt - \frac{c}{(1-d)} \log\left(1-q_0\right)$$

Hence, up to some constant (here $-\frac{c}{(1-d)}\log\left(1-q_0\right)$), our optimal control problem collapses to

$$\max_{0\leq\gamma\leq 1} \int_0^\infty e^{-\rho t} R(q_t) dt$$

where

$$R(q_t) = W^a q_t + \frac{c\rho}{(1-d)} \log\left(1-q_t\right)$$

(ii) Assume now that $d^a \neq d^b$. This objective function can be written as:

$$\int_0^\infty e^{-\rho t} \left\{W^a q_t + \frac{c\left(d^a-d^b\right) q_t}{\left[q_t\left(1-d^a\right)+\left(1-q_t\right)\left(1-d^b\right)\right]} - \frac{c\rho}{\left(1-d^a\right)} \log\left[\frac{1-d^b-q_t\left(d^a-d^b\right)}{1-q_t}\right]\right\} dt$$
$$-\frac{c}{\left(1-d^a\right)} \log\left[\frac{1-q_0}{1-d^b-q_0\left(d^a-d^b\right)}\right]$$

Hence, up to some constant (here $-\frac{c}{(1-d^a)}\log\left[\frac{1-q_0}{1-d^b-q_0\left(d^a-d^b\right)}\right]$), our optimal control problem collapses to

$$\max_{0\leq\gamma\leq 1} \int_0^\infty e^{-\rho t} R(q_t) dt$$

where

$$R(q_t) = W^a q_t + \frac{c\left(d^a-d^b\right) q_t}{\left[q_t\left(1-d^a\right)+\left(1-q_t\right)\left(1-d^b\right)\right]} - \frac{c\rho}{\left(1-d^a\right)} \log\left[\frac{1-d^b-q_t\left(d^a-d^b\right)}{1-q_t}\right]$$

This completes the proof of this lemma. ∎

D.1.2 Equilibrium and dynamics

Let us first consider the case when $d^a = d^b = d$. The objective function $R(q)$ is depicted in Figure D1 when $W^a > c\rho/(1-d)$. It reaches a maximum at $q^* = 1 - c\rho/[W^a(1-d)]$ so that, when $q < q^*$, $R(q)$ is increasing while, when $q > q^*$, $R(q)$ is decreasing. When $W^a < c\rho/(1-d)$, then the function $R(q)$ is always decreasing.

[*Insert Figure D1 here*]

The dynamics is relatively simple in that case. Notice, first, that the cultural dynamics can be written as:

$$\begin{aligned}\overset{\bullet}{q}_t &= (1-q_t)(1-d)[\gamma(1-q_t)+q_t] - q_t(1-q_t)(1-d)(1-\gamma) \quad \text{for } q_t \in]\epsilon, 1-\epsilon[\\ &= (1-d)(1-q_t)\gamma\end{aligned}$$

Hence for given initial conditions $q(0) \in]0,1[$, the fraction of q_t increases in the population as long as $\gamma > 0$. Conversely, given that parental socialization is symmetric, the population stays at any value of q where it is as soon as $\gamma = 0$.

Now, when $W^a < c\rho/(1-d)$, and given some initial conditions $q(0) \in]0,1[$, it is optimal for the leader to remain inactive at all time so that $\gamma^* = 0$. The population stays therefore at this initial value $q(0)$. When $W^a > c\rho/(1-d)$, $q(0) > q^*$, given that γ can only be positive and trigger a further increase in the fraction q, the optimal strategy of the leader is again to remain inactive with $\gamma^* = 0$ and to stay at this initial value $q(0)$. When, however, $q(0) < q^*$, then it pays for the leader to be active up to the point when she reaches q^*. Hence, as long as q_t remains below q^*, the leader provides full biased cultural influence at $\gamma = 1$. Then at some finite time T, the dynamics of q_t imply that the population reaches the targeted fraction q^*. At this time T, the leader stops her cultural influence and chooses to remain inactive at $\gamma^* = 0$. Indeed, when $d^a = d^b = d$, the parent's effort of each type of families is the same. Thus, for $W^a < c\rho/(1-d)$, the gain for the leader to act on γ is too small and therefore she stays inactive ($\gamma^* = 0$). The time pattern of cultural influence for this case is illustrated in Figure D2.

[*Insert Figure D2 here*]

Consider now the case when $d^a \neq d^b$. The study of the shape of the function $R(q_t)$ now depends on whether parental socialization rates of the leader's type (i.e. type a) is large enough compared to those of parents of the other type b. Specifically, we have the following result:

Lemma D2: *Assume that $\Delta d = d^a - d^b \neq 0$ and that the leader is sufficiently patient, i.e. ρ is less than $1 - d^b$.*

(i) When $\Delta d > -\frac{1-d^b}{2}$*, we have:*

(i1) If $\frac{W^a}{c} > \frac{\rho-\Delta d}{1-d^b}$*,the function* $R(q)$ *has a unique maximum at* $q = q_H \in (0,1)$*;*

(i2) If $\frac{W^a}{c} < \frac{\rho-\Delta d}{1-d^b}$*, the function* $R(q)$ *is decreasing for all* $q \in [0,1]$*.*

(ii) When $\Delta d < -\frac{1-d^b}{2}$*, there exists a threshold* $\alpha < \frac{\rho-\Delta d}{1-d^b}$ *such that:*

(ii1) If $\frac{W^a}{c} > \frac{\rho-\Delta d}{1-d^b}$*,the function* $R(q)$ *has a unique maximum at* $q = q_H \in (0,1)$*;*

(ii1) If $\alpha < \frac{W^a}{c} < \frac{\rho-\Delta d}{1-d^b}$*, there exist* q_L *and* q_H *such that the function* $R(q)$ *is decreasing for* $q \in [0,q_L] \cup [q_H,1]$ *and increasing for* $q \in [q_L;q_H]$*;*

(ii3) If $\frac{W^a}{c} < \alpha$*, the function* $R(q)$ *is decreasing in* $q \in [0,1]$*.*

Proof of Lemma D2: The function $R(q_t)$ now takes the following form:

$$R(q_t) = W^a q_t + \frac{c\left(d^a - d^b\right) q_t}{\left[q_t\left(1-d^a\right) + \left(1-q_t\right)\left(1-d^b\right)\right]} - \frac{c\rho}{\left(1-d^a\right)} \log\left[\frac{1-d^b-q_t\left(d^a-d^b\right)}{1-q_t}\right]$$

and differentiation provides

$$R'(q_t) = W^a + \frac{c\left(d^a-d^b\right)\left(1-d^b\right)}{\left[1-d^b-q_t\left(d^a-d^b\right)\right]^2} - \frac{c\rho}{\left[1-d^b-q_t\left(d^a-d^b\right)\right]\left[1-q_t\right]}$$

Now denote $\Lambda(q) = R'(q_t)\left[1-d^b-q_t\left(d^a-d^b\right)\right]^2(1-q_t)$ and pose $\Delta d = \left(d^a - d^b\right)$. Then $\Lambda(q)$ writes as $\Lambda(q) = W^a\Phi(q) - c\Upsilon(q)$ with

$$\begin{aligned}\Phi(q) &= \left[1-d^b-q\Delta d\right]^2(1-q)\\ \Upsilon(q) &= \rho\left(1-d^b-q\Delta d\right) - \Delta d(1-d^b)(1-q)\end{aligned}$$

The function $\Upsilon(q)$ is a linear function of q with $\Upsilon(0) = (\rho-\Delta d)(1-d^b)$ and $\Upsilon(1) = \rho(1-d^a) > 0$. Its slope is $\Delta d(1-d^b-\rho)$. We assume that the discount is small enough that ρ is less than $1-d^b$. Therefore the function $\Upsilon(q)$ is increasing in q if and only if $\Delta d > 0$

The function $\Phi(q) = \left[1-d^b-q\Delta d\right]^2(1-q)$ is such that

$$\begin{aligned}\Phi'(q) &= -2\Delta d\left[1-d^b-q\Delta d\right](1-q) - \left[1-d^b-q\Delta d\right]^2\\ &= -\left[1-d^b-q\Delta d\right]\left[2\Delta d(1-q)+1-d^b-q\Delta d\right]\\ &= -\left[1-d^b-q\Delta d\right]\left[1-d^b+2\Delta d-3\Delta dq\right]\end{aligned}$$

Therefore the function $\Phi(q)$ is decreasing in q when $\Delta d > 0$. It is also decreasing in q when $\Delta d < 0$ and $1-d^b+2\Delta d > 0$. On the other hand when $1-d^b+2\Delta d < 0$ and $\Delta d < 0$, then $\Phi(q)$ is increasing in q if and only if $q \in \left[0, q^M\right]$ with $q^M = \frac{2}{3} + \frac{1-d^b}{3\Delta d}$. Moreover $\Phi(0) = \left[1-d^b\right]^2 > 0$, $\Phi(1) = 0$.

(i) Assume first that $\Delta d > 0$. Then when $\frac{W^a}{c} > \frac{\rho-\Delta d)}{1-d^b}$, there exists a unique q_H such that $\Lambda(q)$ is positive if and only if $q \in [0,q_H]$ and the function $R(q)$ has a unique maximum at $q = q_H$. When $\frac{W^a}{c} \leq \frac{\rho - \Delta d}{1-d^b}$, $\Lambda(q)$ is always negative for all q and the function $R(q)$ is decreasing in $q \in [0,1]$.

(ii) Suppose now that $\Delta d < 0.$ and $1 - d^b + 2\Delta d > 0$. Then $\Phi(q)$ is decreasing in q. Moreover it is easy to see that $\Phi''(q) > 0$. $\Upsilon(q)$ is also linear decreasing in q. Again it is easy to see that when $\frac{W^a}{c} > \frac{\rho - \Delta d}{1-d^b}$, there exists a unique q_H such that $\Lambda(q)$ is positive if and only if $q \in [0, q_H]$ and the function $R(q)$ has a unique maximum at $q = q_H$. When $\frac{W^a}{c} \leq \frac{\rho - \Delta d}{1-d^b}$, $\Lambda(q)$ is always negative for all q and the function $R(q)$ is decreasing in $q \in [0,1]$.

(iii) Suppose now that $\Delta d < 0.$ and $1 - d^b + 2\Delta d < 0$. $\Phi(q)$ is increasing in q if and only if $q \in \left[0, q^M\right]$. Denote $\Phi^M = \Phi(q^M)$ the maximum of $\Phi(q)$. $\Upsilon(q)$ is linear decreasing in q. . When $\frac{W^a}{c} > \frac{\rho - \Delta d}{1-d^b}$, it is easy to see that as before there exists a unique q_H such that $\Lambda(q)$ is positive if and only if $q \in [0, q_H]$ and the function $R(q)$ has a unique maximum at $q = q_H$. When $\frac{W^a}{c} < \frac{\rho - \Delta d}{1-d^b}$, the function $\Phi(q)$ is decreasing in q. Moreover it is easy to see that $\Phi''(q) < 0$ if and only if $q \leq q_C = \frac{1}{3} + \frac{2(1-d^b)}{3\Delta d} > q^M$.

The shape of the functions $\Phi(q)$ and $\frac{c}{W^a}\Upsilon(q)$ is depicted in Figure D3 (when $\frac{W^a}{c} > \frac{\rho - \Delta d}{1-d^b}$) and Figure D4 ($\frac{W^a}{c} \leq \frac{\rho - \Delta d}{1-d^b}$).

[*Insert Figures D*3 *and D*4 *here*]

As can be seen at $\frac{W^a}{c} = \frac{\rho - \Delta d}{1-d^b}$ one has $\frac{c}{W^a}\Upsilon(0) = \Phi(0)$ while $\frac{c}{W^a}\Upsilon(1) > 0 = \Phi(1)$. There is therefore a unique intersection at a point q_H . By continuity , a decrease of $\frac{W^a}{c}$ from $\frac{\rho - \Delta d}{1-d^b}$ leads to an upward shift of the function $\frac{c}{W^a}\Upsilon(q)$. The function $\frac{c}{W^a}\Upsilon(q)$ crosses then two times the function $\Phi(q)$ at the two points $q_L\left(\frac{W^a}{c}\right)$ and $q_H\left(\frac{W^a}{c}\right)$ which are respectively decreasing and increasing functions of $\frac{W^a}{c}$). From this it follows that the function $\Lambda(q)$ is positive for $q \in \left[q_L\left(\frac{W^a}{c}\right); q_H\left(\frac{W^a}{c}\right)\right]$ and negative for $q \in \left[0, q_L\left(\frac{W^a}{c}\right)\right] \cup \left[q_H\left(\frac{W^a}{c}\right), 1\right]$. It follows immediately that in such case the function $R(q)$ has a unique minimum at $q = q_L$ and a local maximum at $q = q_H$.

Keeping on decreasing further the parameter$\frac{W^a}{c}$ leads the two points q_L and q_H to converge by continuity to the tangent point $q^T \in \left(q^M; q^C\right)$ between $\frac{c}{W^a}\Upsilon(q)$ and $\Phi(q)$. When $\frac{W^a}{c}$ is further decreased below the value α such that $q^T = q_L(\alpha)$, the curve $\frac{c}{W^a}\Upsilon(q)$ is always above the curve $\Phi(q)$. From this $\Lambda(q)$ is always negative for all q and the function $R(q)$ is decreasing in $q \in [0,1]$.

This completes the proof of this lemma. ▮

Again assume that without leader intervention, the dynamics of culture have converged towards their steady states $q(0) = \epsilon$ when $\Delta d < 0$, or $q(0) = 1 - \epsilon$ when $\Delta d > 0$. Then we have the following proposition:

Proposition D1: *Assume that* $\Delta d = d^a - d^b \neq 0$ *and that the leader is sufficiently patient, i.e.* ρ *is less than* $1 - d^b$. *For any* ϵ *small enough, we have:*

(i) When $\Delta d > 0$ and $q(0) = 1 - \epsilon$, then the leader never socializes, i.e. $\gamma^* = 0$ and the cultural dynamics stay at $q(0) = 1 - \epsilon$.

(ii) When $\Delta d < 0$ and $q(0) = \epsilon$, then we have:

(ii1) If $\frac{W^a}{c} > \frac{\rho - \Delta d}{1 - d^b}$, the leader chooses to fully socialize $\gamma^* = 1$ up to to some finite time T for which $q(T) = q_H$ and then, for $t > T$, to socialize at the interior rate

$$\gamma^* = \frac{-q_H \Delta d}{(1 - d^b) - q_H \Delta d}$$

The cultural dynamics then stay at the interior steady state q_H.

(ii2) If $\frac{W^a}{c} < \frac{\rho - \Delta d}{1 - d^b}$, the leader never socializes, $\gamma^* = 0$, and the cultural dynamics stay at $q(0) = \epsilon$.

Proof of Proposition D1: We have seen that, for $\Delta d = d^a - d^b \neq 0$ and for the leader sufficiently patient (i.e. ρ less than $1 - d^b$), the shape of the function $R(q)$ is given by Lemma D2. Then it is easy to have the following result:

(i) Consider first that $\Delta d > 0$ and $q(0) = 1 - \epsilon$. When $\frac{W^a}{c} > \frac{\rho - \Delta d}{1 - d^b}$ and for ϵ small enough $1 - \epsilon > q_H$. Therefore any positive value of γ would only increase further q and reduce the value of the function $R(q)$. Therefore the leader should stay at $q(0) = 1 - \epsilon$ and choose $\gamma^* = 0$. When $\frac{W^a}{c} < \frac{\rho - \Delta d}{1 - d^b}$, the function $R(q)$ is decreasing for all values of q. Therefore again the leader should stay at $q(0) = 1 - \epsilon$ and choose $\gamma^* = 0$.

(ii) Consider now that $\Delta d < 0$ and $q(0) = \epsilon$. then:

$(ii1)$ when $\frac{W^a}{c} > \frac{\rho - \Delta d}{1 - d^b}$, the function $R(q)$. has again a global maximum at $q = q_H$. As a consequence given that $\epsilon < q_H$, the leader chooses optimally his MRAP towards q_H with full socialization $\gamma^* = 1$ up to some finite time T at which $q(T) = q_H$. Then for $t > T$, the leader maintains the cultural dynamics at $q = q_H$ with an interior socialization rate γ such that $\dot{q} = 0$, namely:

$$\gamma^* = \frac{-q_H \Delta d}{(1 - d^b) - q_H \Delta d}$$

The cultural dynamics then stay at the interior steady state q_H.

$(ii2)$ When $\frac{W^a}{c} \in]\alpha, \frac{\rho - \Delta d}{1 - d^b}[$, there exists q_L and q_H such that the function $R(q)$ is decreasing for $q \in [0, q_L] \cup [q_H, 1]$ and increasing for $q \in [q_L; q_H]$. For ϵ small enough (i.e. $\epsilon < q_L$) , the intertemporal value $V(\epsilon) = \frac{R(\epsilon)}{\rho}$ is larger than $\frac{R(q_H)}{\rho}$, and therefore is larger than any other dynamic path that converges monotonically from $q(0)$ to q_H. It is then optimal for the leader to stay at $q = \epsilon$ with $\gamma^* = 0$.

$(ii3)$ when $\frac{W^a}{c} < \alpha$, the function $R(q)$ is decreasing in $q \in [0, 1]$. Hence again the leader never socializes, $\gamma^* = 0$. and the cultural dynamics stay at $q(0) = \epsilon$. ∎

Intuitively, Proposition D1 provides very similar results to the case where the cultural leader is not directly participating into the socialization process of her trait (benchmark model). Specifically, when families of type a are more successful in socializing their offsprings than families of type b, the cultural leader does not enter into the process of socialization as her action is a pure substitute to the family socialization from members of her own community. More interestingly, when family socialization by individuals of trait a is lower than that of individuals of trait b, then a cultural leader may enter actively into direct socialization when the relative value of the cultural rent W^a/c she derives from this socialization process is above a certain threshold (here $(\rho - \Delta d) / \left(1 - d^b\right)$). Such leader will socialize to compensate for the lack of family socialization to her own trait and, by doing so, will induce a positive long-term fraction q_H of members of her own type. As in the previous section, the direct socialization effort of the leader will temporarily overshoot her long-run value. Indeed, for a finite period of time, the leader will provide a maximum socialization of $\gamma = 1$. Then, after reaching the optimal population share of q_H, the leader will reduce the socialization effort to an intermediate value of $\gamma^* < 1$ such that the system stays at q_H. Note that the leader is more likely to be active the more patient she is (i.e. the smaller is ρ), the smaller d^b (the family effort of the other group b) and the larger the difference $-\Delta d = d^b - d^a$ between family socialization of types b and a. Obviously, the emergence of a cultural leader for group a is also more likely to happen when the cost of direct socialization c is small.

Our analysis thus suggests that, in the context of religious radical leaders (group a), they are likely to be active when, in the community, family socialization to the mainstream model is weak (d^b small) but still larger than family socialization for radicalism ($\Delta d < 0$). In such situation, when the cost of direct socialization is low (due, for example, to easier access to communication technologies), cultural leaders, supporting traits that would have otherwise disappeared (such as Muslim radicalization), have larger incentives to become active to countervail the natural tendency of the long-run assimilation of the ethnic population to the mainstream host culture.

More generally, all these extensions of the benchmark model highlight the idea that, independently of the objective function of cultural leaders, the nature of the intervention process of the leader or the fact that the public good can be imperfectly excludable to outgroups, the qualitative nature of the dynamics of cultural integration crucially depend on the cultural substituability and internalization (complementarity) effects between centralized and decentralized channels of cultural transmission inside communities. The existence of these two types of effects then typically generate non-monotonic size effects, threshold and hysteresis effects, which are crucial determinants of these dynamics.

D.2 Competition between cultural leaders

In this section, we extend our framework of direct socialization of the leader to discuss the issue of *competition between cultural leaders*, for example, between a religious leader and the host-country institution. The transmission process is now as follows. First, parents of type $i \in \{a, b\}$ directly transmit their trait with probability d^i. To have a tractable setting, we assume that family socialization d^i is exogenous. Second, cultural influence *outside* the family depends on two types of role models. As before, there are *passive role models* with whom naive children can be randomly matched. With probability $1-\gamma$, the child is matched with a passive role model randomly chosen in the society. In such case, she adopts trait a with probability $q^a = q$ and trait b with probability $q^b = 1 - q$. There are also *specific role models* (i.e. *cultural leaders*) who do interact strategically by choosing their socialization efforts, *directly* influencing cultural evolution in a way that favors their interests. It is clear that these two leaders will compete since their objectives are different, even opposite. To be more precise, with probability γ, the child will be exposed to a cultural leader of type $i \neq j$. She then gets socialized to trait i with probability $\Pi^i(\phi^i, \phi^j)$ and, with probability $\Pi^j(\phi^j, \phi^i) = 1 - \Pi^i(\phi^i, \phi^j)$, she adopts the other trait $j \neq i$, where ϕ^i and ϕ^j are the socialization efforts of the leaders of type i and type j, respectively. Observe that γ reflects the relative importance of cultural leaders in the transmission process as compared to the rest of society. This is a technological parameter that may also reflect the importance of information technologies or the degree of centrality of the network of cultural socialization in society. For simplicity, we take a specific well-known contest function (Hirshleifer, 1989; Skaperdas, 1996; Konrad, 2009) given by:

$$\Pi^a(\phi^a, \phi^b) = 1 - \Pi^b(\phi^a, \phi^b) = \frac{\phi^A}{\phi^A + \phi^B} \tag{D.4}$$

This formulation captures the fact that cultural leaders are in competition with each other for cultural transmission. Therefore, for all $i, j \in \{a, b\}$, the probability that a child from a family with trait i is socialized to trait j is now given by:

$$P^{ii}\left(d^i, \phi^i, \phi^j, \gamma, q^i\right) = d^i + (1 - d^i)\gamma\Pi^i(\phi^i, \phi^j) + (1 - \gamma)q^i$$

$$P^{ij}\left(d^i, \phi^i, \phi^j, \gamma, q^i\right) = (1 - d^i)\gamma\Pi^j(\phi^j, \phi^i) + (1 - \gamma)(1 - q^i)$$

The dynamics of the fraction of the population with trait i, can then be straightforwardly written as:

$$q^i_{t+1} = q^i_t P^{ii}\left(d^i, \phi^i, \phi^j, \gamma, q^i\right) + (1 - q^i_t)P^{ji}\left(d^j, \phi^i, \phi^j, \gamma, q^i\right)$$

Thus, taking as before the notation $q^a = 1 - q^b = q$ and passing to continuous time dynamics, we obtain:

$$\overset{\bullet}{q} = q(1-q)(1-\gamma)\left(d^a - d^b\right) + \gamma\left[(1-q)(1-d^b)\Pi^a - q(1-d^a)(1-P^a)\right] \tag{D.5}$$

where $\Pi^a \equiv \Pi^a(\phi^A, \phi^B)$.

We assume that cultural leaders enjoy the rents associated with the number of individuals who share their cultural traits in the population and therefore are ready to spend resources to affect the process of cultural evolution in society. Specifically, assume that the utility of a cultural leader of group $i \in \{a, b\}$ is given by:

$$V^i = \int_0^\infty e^{-\rho^i t}\left(W^i q^i - c^i \phi^i\right) dt$$

where the leader's cultural rents $W^a q$ (resp. $W^b(1-q)$) of type a (resp. b) increases with the size of her own group $q^a = q$ (resp. $q^b = 1-q$) and $c^a\phi^a$ (resp. $c^b\phi^b$) is the linear resource cost to compete for cultural socialization in the contest between leaders. This setting describes then a *dynamic differential game* of cultural influence between the two leaders. As it is well-known, various equilibrium concepts can be used to analyze such games. In the following, we characterize the *open-loop Nash equilibrium* concept. Denote by $\phi^a(.)$ and $\phi^b(.) \in \mathbb{R}_+$ some admissible socialization efforts of the leaders a and b, and by $V^a\left(\phi^a(.), \phi^b(.)\right)$ and $V^b\left(\phi^a(.), \phi^b(.)\right)$ the intertemporal values associated to these socialization efforts, i.e.,

$$V^a\left(\phi^a(.), \phi^b(.)\right) = \int_0^\infty e^{-\rho^a t}\left[W^a q(t, \phi^a, \phi^b) - c^a \phi^a(t)\right] dt$$

$$V^b\left(\phi^a(.), \phi^b(.)\right) = \int_0^\infty e^{-\rho^b t}\left[W^b(1 - q(t, \phi^a, \phi^b)) - c^b \phi^b(t)\right] dt$$

where $q(t, \phi^a, \phi^b)$ is the trajectory starting at q_0 and satisfying the following differential equation:

$$\overset{\bullet}{q}_t = q_t(1-q_t)(1-\gamma)\left[d^a - d^b\right] + \gamma\left[(1-q_t)(1-d^b)\frac{\phi_t^a}{\phi_t^a + \phi_t^b} - q_t(1-d^a)\frac{\phi_t^b}{\phi_t^a + \phi_t^b}\right] \tag{D.6}$$

Definition D1: *An open-loop Nash equilibrium* $(\phi_t^{a*}, \phi_t^{b*})$ *is characterized by the following conditions:*

$$V^a\left(\phi^{a*}(.), \phi^{b*}(.)\right) \geq V^a\left(\phi^a(.), \phi^{b*}(.)\right) \text{ for all admissible } \phi^a(.)$$

$$V^b\left(\phi^{a*}(.), \phi^{b*}(.)\right) \geq V^b\left(\phi^{a*}(.), \phi^{b}(.)\right) \text{ for all admissible } \phi^b(.)$$

In the Online Appendix D.3, we characterize the solutions of the open-loop Nash equilibrium $(\phi_t^{a*}, \phi_t^{b*})$ and show that a complete characterization of the equilibrium trajectory and socialization efforts of the two leaders is intractable. As a result, to obtain analytical results, we assume that $d^a = d^b = d$, which means that the family (vertical) socialization is neutral in terms of cultural evolution so that both parents (of types a and b) provide the same constant socialization effort. In other words, we shut down the vertical socialization channel and focus on the oblique one by analyzing the influence of each type of leader on socialization when they compete with each other. In such a case, the conditions given in the Online Appendix D.3 collapse to much simpler conditions, which are:

$$\begin{array}{l} \lambda^a\gamma(1-d)\frac{\phi^{b*}}{\left(\phi^{a*}+\phi^{b*}\right)^2} = c^a \\ -\lambda^b\gamma(1-d)\frac{\phi^{a*}}{\left(\phi^{a*}+\phi^{b*}\right)^2} = c^b \\ \overset{\bullet}{\lambda^a} - \rho^a\lambda^a = -\left[W^a - \gamma\lambda^a(1-d)\right] \quad \text{and} \quad \lim_{t\to\infty}\lambda_t^a e^{-\rho^a t} = 0 \\ \overset{\bullet}{\lambda^b} - \rho^b\lambda^b = -\left[-W^b - \gamma\lambda^b(1-d)\right] \quad \text{and} \quad \lim_{t\to\infty}\lambda_t^b e^{-\rho^b t} = 0 \\ \overset{\bullet}{q_t} = \gamma(1-d)\left(\frac{\phi^{a*}}{\phi^{a*}+\phi^{b*}} - q_t\right) \end{array} \tag{D.7}$$

where λ^a and λ^b are the associated adjoint variables to each leader. It follows immediately that the adjoint equilibrium variables jump immediately to their long-run values, given by:

$$\lambda^a = \frac{W^a}{\rho^a + \gamma(1-d)}, \; \lambda^b = \frac{-W^b}{\rho^b + \gamma(1-d)}$$

while the equilibrium socialization efforts of the two leaders are constant and equal to:

$$\frac{\lambda^a}{-\lambda^b}\frac{\phi^{b*}}{\phi^{a*}} = \frac{c^a}{c^b}$$

or equivalently

$$\frac{\phi^{a*}}{\phi^{b*}} = \frac{c^b}{c^a}\frac{W^a}{W^b}\left[\frac{\rho^b + \gamma(1-d)}{\rho^a + \gamma(1-d)}\right] \tag{D.8}$$

This implies that the ratio of socialization efforts of the two cultural leaders will depend on the relative cost and benefit of providing socialization effort, on the relative discount rate or degree of patience and on γ, the relative importance of cultural leaders in the transmission process as compared to the rest of society. On can verify that the sign of the impact of γ on ϕ^{a*}/ϕ^{b*} is the same as $\rho^a - \rho^b$. In other words, if the cultural leader of type a is more (less) patient than that of type b, then the higher is the importance of leaders in the society, the higher (lower) is the effort of leader a compared to that of leader b. We can then determine each socialization effort as follows:

$$\phi^{a*} = \frac{W^b c^b (W^a)^2 \left[\rho^b + \gamma(1-d)\right] \gamma(1-d)}{\left\{c^b W^a \left[\rho^b + \gamma(1-d)\right] + c^a W^b \left[\rho^a + \gamma(1-d)\right]\right\}^2} \tag{D.9}$$

$$\phi^{b*} = \frac{W^a c^a \left(W^b\right)^2 \left[\rho^a + \gamma(1-d)\right] \gamma(1-d)}{\left\{c^b W^a \left[\rho^b + \gamma(1-d)\right] + c^a W^b \left[\rho^a + \gamma(1-d)\right]\right\}^2} \tag{D.10}$$

The cultural dynamics (D.6) are now given by (when $d^a = d^b = d$):

$$\overset{\bullet}{q}_t = \gamma(1-d)\left(q^* - q_t\right)$$

where $q^* = \phi^{a*}/\left(\phi^{a*} + \phi^{b*}\right)$ is the long-run value of the fraction of individuals of type a and, using (D.9) and (D.10), is equal to:

$$q^* = \frac{c^b W^a \left[\rho^b + \gamma(1-d)\right]}{c^b W^a \left[\rho^b + \gamma(1-d)\right] + c^a W^b \left[\rho^a + \gamma(1-d)\right]} \tag{D.11}$$

Note that, in this case, the current-value Hamiltonian $H^a(\phi^a, \phi^b, q, \lambda^a)$ (resp. $H^b(\phi^a, \phi^b, q, \lambda^b)$) is the sum of a linear function in q and a strictly concave function in ϕ^a (resp. in ϕ^b) and thus is jointly strictly concave in (ϕ^a, q) (resp. $\left(\phi^b, q\right)$). Hence, given the other leader's optimal socialization rate, the optimal control problem of each cultural leader satisfies the Magasarian sufficiency conditions and the necessary conditions (D.7) are also sufficient to characterize the Nash open-loop equilibrium.

The phase diagram of the cultural evolution process is depicted in Figure D5, which immediately shows that the steady state q^* is unique and stable.

[*Insert Figure D5 here*]

Proposition D2: *The steady state interior equilibrium q^* is increasing in c^b, ρ^b, W^a and decreasing in c^a, ρ^a, W^b. Moreover, the sign of the impact of $\gamma(1-d)$ on q^* is the same as that of $\rho^a - \rho^b$, i.e.*

$$\frac{\partial q^*}{\partial\left[\gamma(1-d)\right]} \gtreqless 0 \Leftrightarrow \rho^a \gtreqless \rho^b.$$

The steady-state equilibrium fraction of people of type a decreases with the cost and the discount factor ρ^a of the leader of type a and increases with the cost and the discount factor ρ^b of the leader of type b. We have the opposite result for the benefit of leaders of type

$i \in \{a, b\}$. A more surprising result is the impact of γ and d on q^*, which depends on the relative discount rate (or degree of impatience) of the two leaders. Indeed, if the cultural leader of type a is less patient than the leader of type b (i.e. $\rho^a > \rho^b$), then the higher is the importance of leaders in the society (γ), the higher is the fraction of individual of type a in the long-run equilibrium. The intuition for this result is as follow. When γ increases (because of, say, technical progress in communication technologies), the competition between leaders is intensified and both socialization efforts ϕ^a and ϕ^b increase. However, the more patient leader (of type b) initially exerts a higher socialization effort ϕ^b than the less patient leader of type a. Given that the contest function features marginal decreasing returns in own influence effort, the more patient leader b is therefore increasing her socialization effort by less than the more impatient leader a, in response to a higher value of γ. As a consequence, the relative socialization efforts ratio ϕ^a/ϕ^b tends to increase, leading to a higher relative socialization success of leader a compared to leader b. The long run consequence of this is an increase in the steady-steady value q^* of individuals with trait a in the population.

Similarly, if the cultural leader of type a is less (more) patient than that of type b, then the higher is the effort exerted by parents of both types in socializing their kids, i.e. higher d, the lower (higher) is q^*. Again, this is due to the cultural substituability effect between vertical and oblique socialization and its differential impact on the cultural leaders' incentives to socialize depending on their degree of patience. Indeed, when d increases, there is less scope for leader influence and therefore less competition between cultural leaders. Since the less patient leader (say leader a when $\rho^a > \rho^b$) is reducing more strongly her effort than the patient leader (because of the decreasing marginal returns of the contest function and the fact that she already invests less in socialization than the patient leader), then ϕ^{a*}/ϕ^{b*} goes down. As a consequence leader a is relatively less successful at socializing members of the population to trait a and the long-run fraction q^* of individuals of type a is reduced.

References

[1] Hirshleifer, J. (1989), "Conflict and rent-seeking success functions: Ratio vs. difference models of relative success," *Public Choice* 63, 101-112.

[2] Konrad, K.A. (2009), *Strategy and Dynamics in Contests*, New York: Oxford University Press.

[3] Skaperdas, S. (1996), "Contest success functions," *Economic Theory* 7, 283-290.

[4] Verdier, T. and Y. Zenou (2015), "The role of cultural leaders in the transmission of preferences," *Economics Letters* 136, 158-161.

D.3 Characterization of the solutions of the open-loop Nash equilibrium $(\phi_t^{a*}, \phi_t^{b*})$

To characterize the solutions of the open-loop Nash equilibrium $(\phi_t^{a*}, \phi_t^{b*})$, we need to provide the following current-value Hamiltoneans:

$$\begin{aligned} & H^a(\phi^a, \phi^b, q, \lambda^a) \\ = \; & W^a q - c^a \phi^a \\ & + \lambda^a \left[q\,(1-q)\,(1-\gamma)\left(d^a - d^b\right) + \gamma\left((1-q)(1-d^b)\frac{\phi^a}{\phi^a+\phi^b} - q(1-d^a)\frac{\phi^b}{\phi^a+\phi^b}\right)\right] \end{aligned}$$

$$\begin{aligned} & H^b(\phi^a, \phi^b, q, \lambda^b) \\ = \; & W^b (1-q) - c^b \phi^b \\ & + \lambda^b \left[q\,(1-q)\,(1-\gamma)\left(d^a - d^b\right) + \gamma\left((1-q)(1-d^b)\frac{\phi^a}{\phi^a+\phi^b} - q(1-d^a)\frac{\phi^b}{\phi^a+\phi^b}\right)\right] \end{aligned}$$

where λ^a and λ^b are the associated adjoint variables to each leader and the parental socialization efforts take the general form $d^a = d^a(q)$; $d^b = d^b(q)$. An open loop Nash equilibrium $(\phi^{a*}(.), \phi^{b*}(.))$ should then satisfy the following necessary conditions

$$\begin{aligned} \phi^{a*} &= \arg\max_{\phi^a} H^a(\phi^a, \phi^{b*}, q, \lambda^a) \\ \phi^{b*} &= \arg\max_{\phi^b} H^b(\phi^{a*}, \phi^b, q, \lambda^b) \\ \overset{\bullet}{\lambda}{}^a - \rho^a \lambda^a &= -\frac{\partial H^a(\phi^{a*}, \phi^{b*}, q, \lambda^a)}{\partial q} \quad \text{and} \quad \lim_{t\to\infty} \lambda^a(t) e^{-\rho^a t} = 0 \\ \overset{\bullet}{\lambda}{}^b - \rho^b \lambda^b &= -\frac{\partial H^b(\phi^{a*}, \phi^{b*}, q, \lambda^b)}{\partial q} \quad \text{and} \quad \lim_{t\to\infty} \lambda^b(t) e^{-\rho^b t} = 0 \end{aligned}$$

$$\begin{aligned} \overset{\bullet}{q}_t = \; & q_t\,(1-q_t)\,(1-\gamma)\left(d^a(q_t) - d^b(q_t)\right) \\ & + \gamma\left[(1-q_t)(1-d^b(q_t))\frac{\phi_t^{a*}}{\phi_t^{a*}+\phi_t^{b*}} - q_t(1-d^a(q_t))\frac{\phi_t^{b*}}{\phi_t^{a*}+\phi_t^{b*}}\right] \end{aligned}$$

or equivalently,

$$\lambda^a \gamma\left[(1-q)(1-d^b(q)) + q(1-d^a(q))\right]\frac{\phi^{b*}}{\left(\phi^{a*}+\phi^{b*}\right)^2} = c^a$$

$$-\lambda^b \gamma\left[(1-q)(1-d^b(q)) + q(1-d^a(q))\right]\frac{\phi^{a*}}{\left(\phi^{a*}+\phi^{b*}\right)^2} = c^b$$

$$\overset{\bullet}{\lambda}^a - \rho^a \lambda^a$$
$$= -\left\{ \begin{array}{c} W^a + \lambda^a (1-2q)) (1-\gamma) \left(d^a - d^b\right) - \gamma\lambda^a \left[(1-d^b)\frac{\phi^{a*}}{\phi^{a*}+\phi^{b*}} + (1-d^a)\frac{\phi^{b*}}{\phi^{a*}+\phi^{b*}}\right] \\ +\lambda^a \cdot \Theta\left(q, \phi^{a*}, \phi^{b*}\right) \end{array} \right\} -$$
$$\text{and } \lim_{t\to\infty} \lambda_t^a e^{-\rho^a t} = 0$$

$$\overset{\bullet}{\lambda}^b - \rho^b \lambda^b$$
$$= -\left\{ \begin{array}{c} -W^b + \lambda^b (1-2q)) (1-\gamma) \left(d^a - d^b\right) - \gamma\lambda^b \left[(1-d^b)\frac{\phi^{a*}}{\phi^{a*}+\phi^{b*}} + (1-d^a)\frac{\phi^{b*}}{\phi^{a*}+\phi^{b*}}\right] \\ +\lambda^b \cdot \Theta\left(q, \phi^{a*}, \phi^{b*}\right) \end{array} \right\}$$
$$\text{and } \lim_{t\to\infty} \lambda_t^b e^{-\rho^b t} = 0$$
$$\overset{\bullet}{q}_t = q_t (1-q_t)(1-\gamma)\left(d^a - d^b\right) + \gamma\left[(1-q_t)(1-d^b)\frac{\phi_t^{a*}}{\phi_t^{a*}+\phi_t^{b*}} - q_t(1-d^a)\frac{\phi_t^{b*}}{\phi_t^{a*}+\phi_t^{b*}}\right]$$

denoting $\Theta\left(q, \phi^a, \phi^b\right)$, the following function :

$$\Theta\left(q, \phi^a, \phi^b\right) = -(1-q)d^{b\prime}(q))\left[\gamma\frac{\phi^a}{\phi^a+\phi^b} + (1-\gamma)\,q\right] + qd^{a\prime}(q)\left[1 - \gamma\frac{\phi^a}{\phi^a+\phi^b} - q\,(1-\gamma)\right]$$

It should be clear that a complete characterization of the equilibrium trajectory and socialization efforts of the two leaders is intractable in this general case.

Figure B1(a): $W^a < \Phi(\hat{q})$

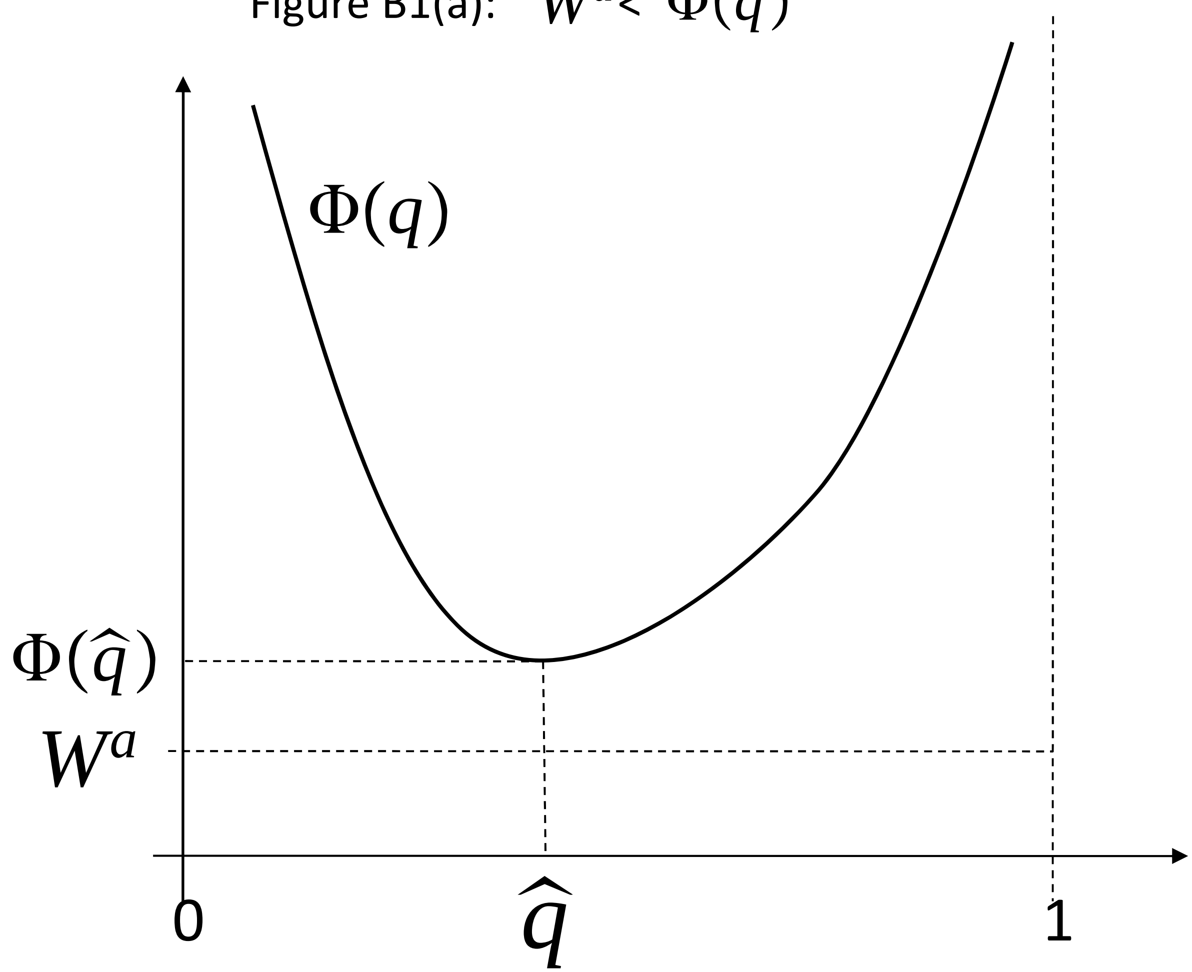

Figure B1(b): $W^a > \Phi(\hat{q})$

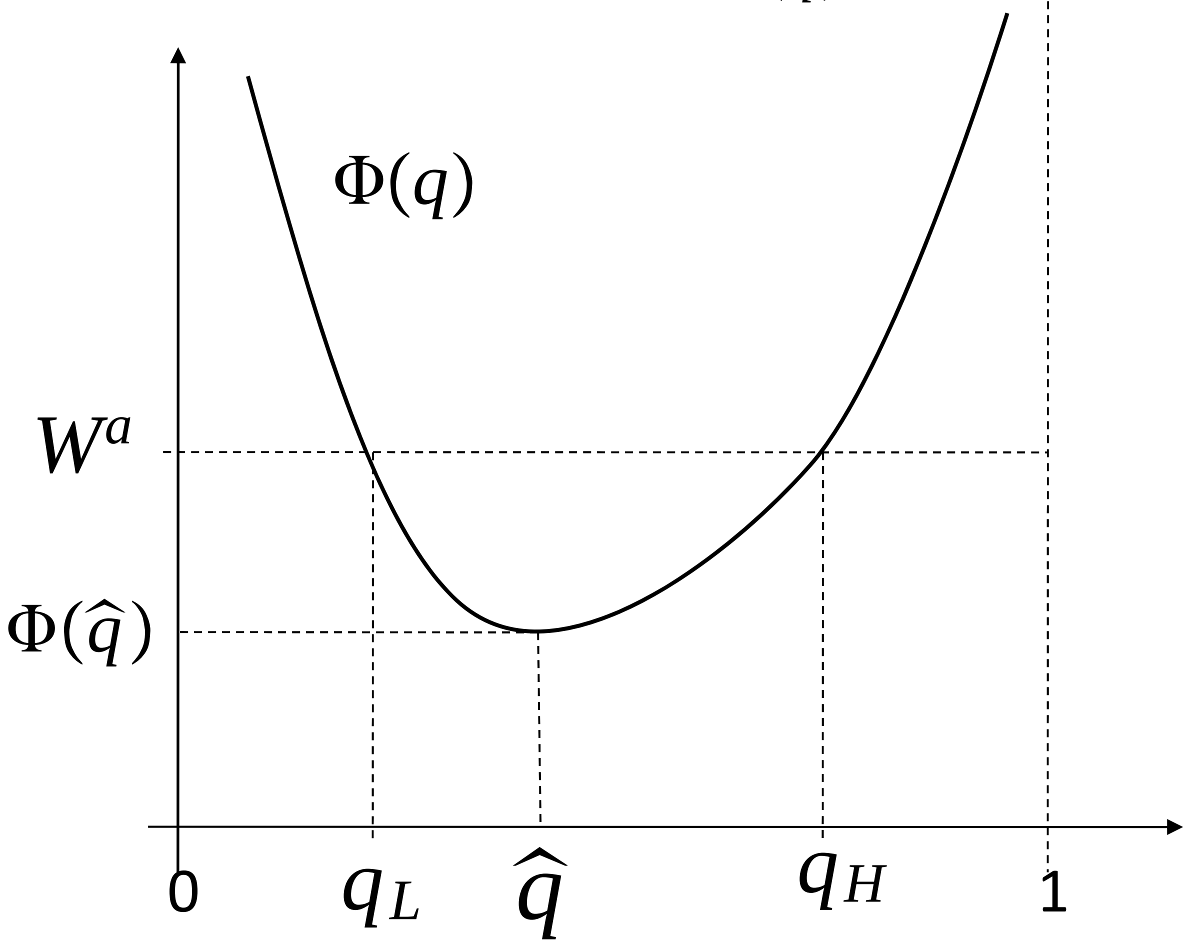

Figure B2: The function $R(q)$ with public good spillovers

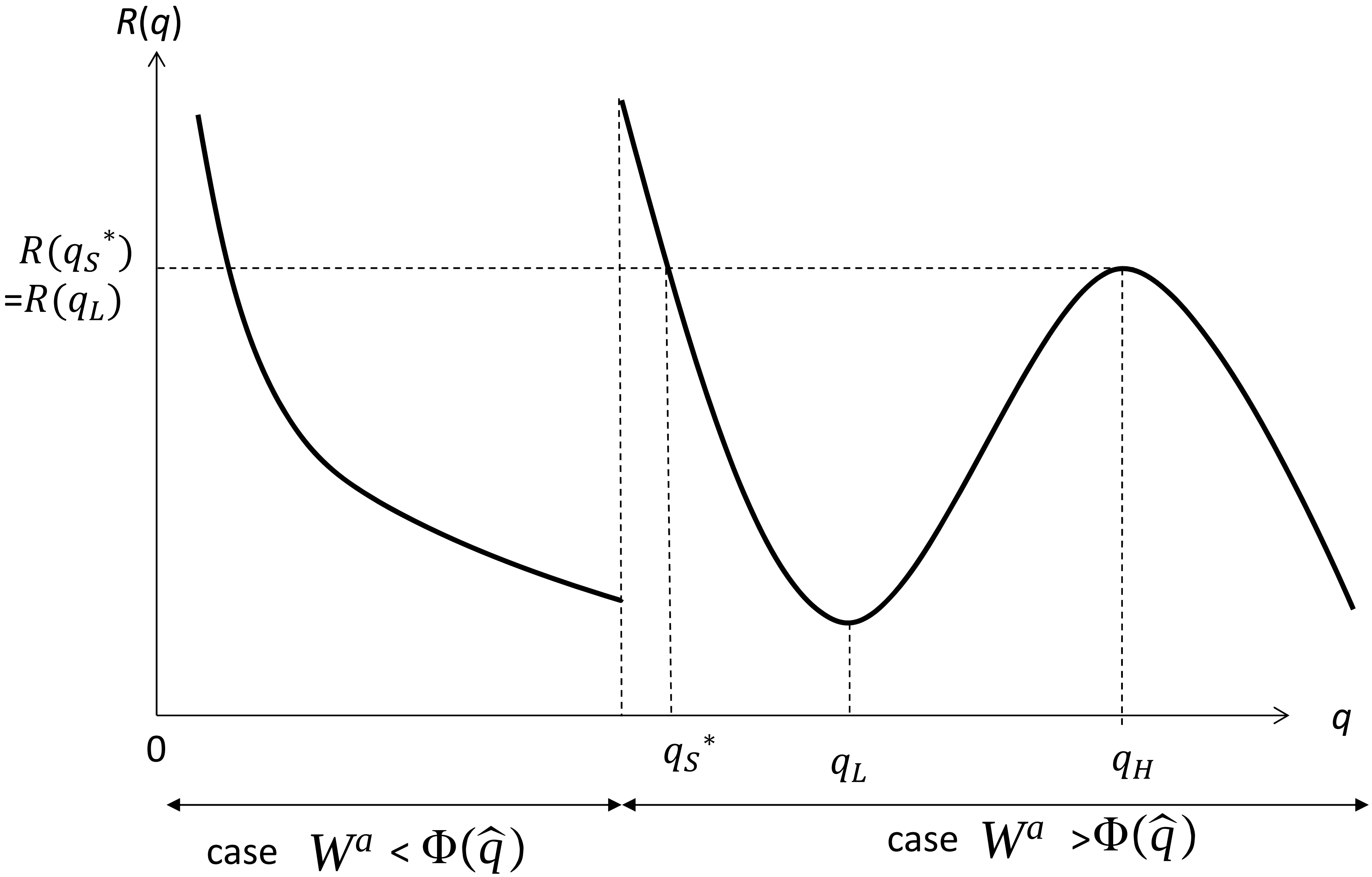

Figure C1: The function $R(q)$

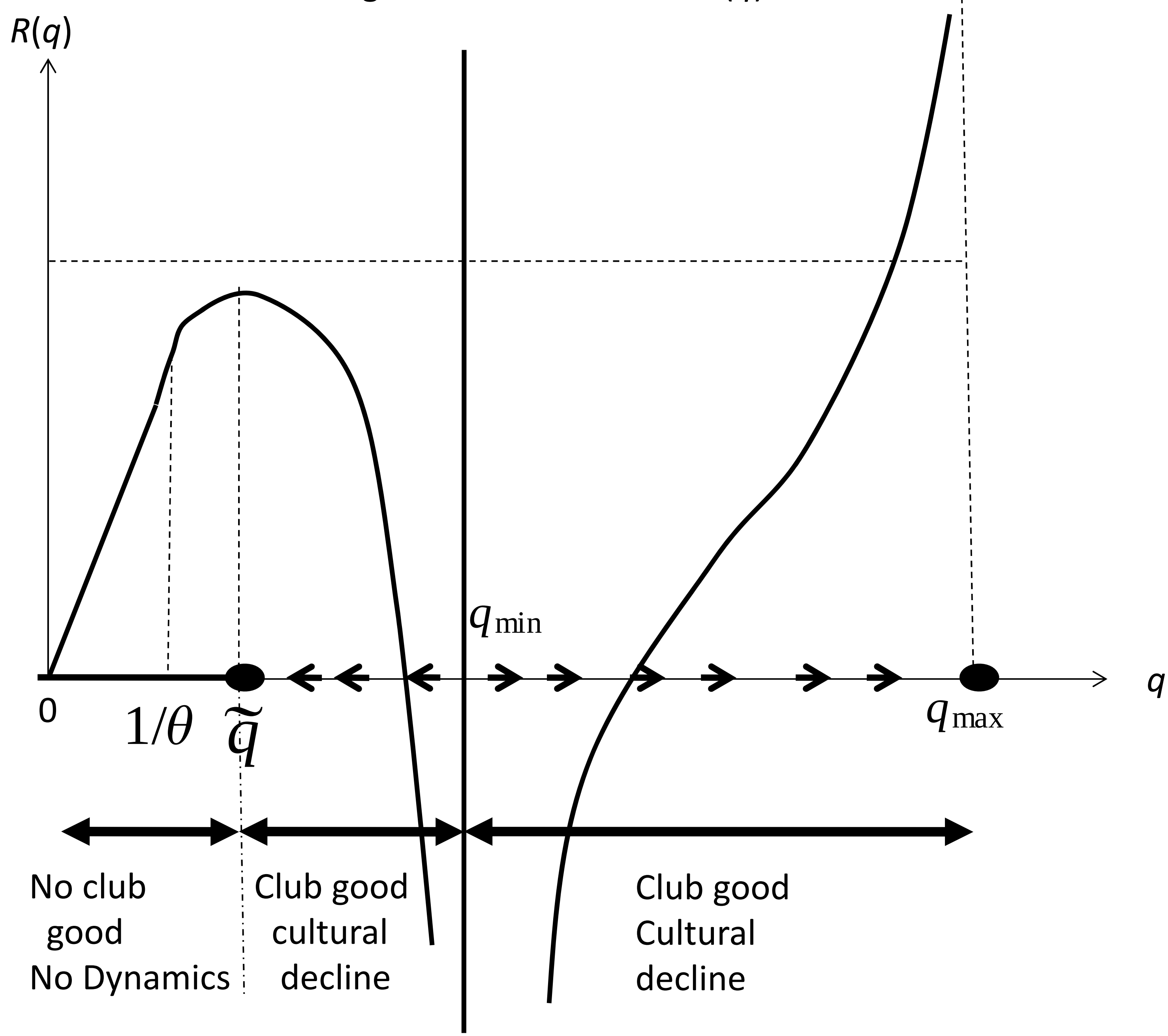

Figure D1: The function $R(q)$ when the leader chooses γ when $W^a > \frac{c\rho}{1-d}$

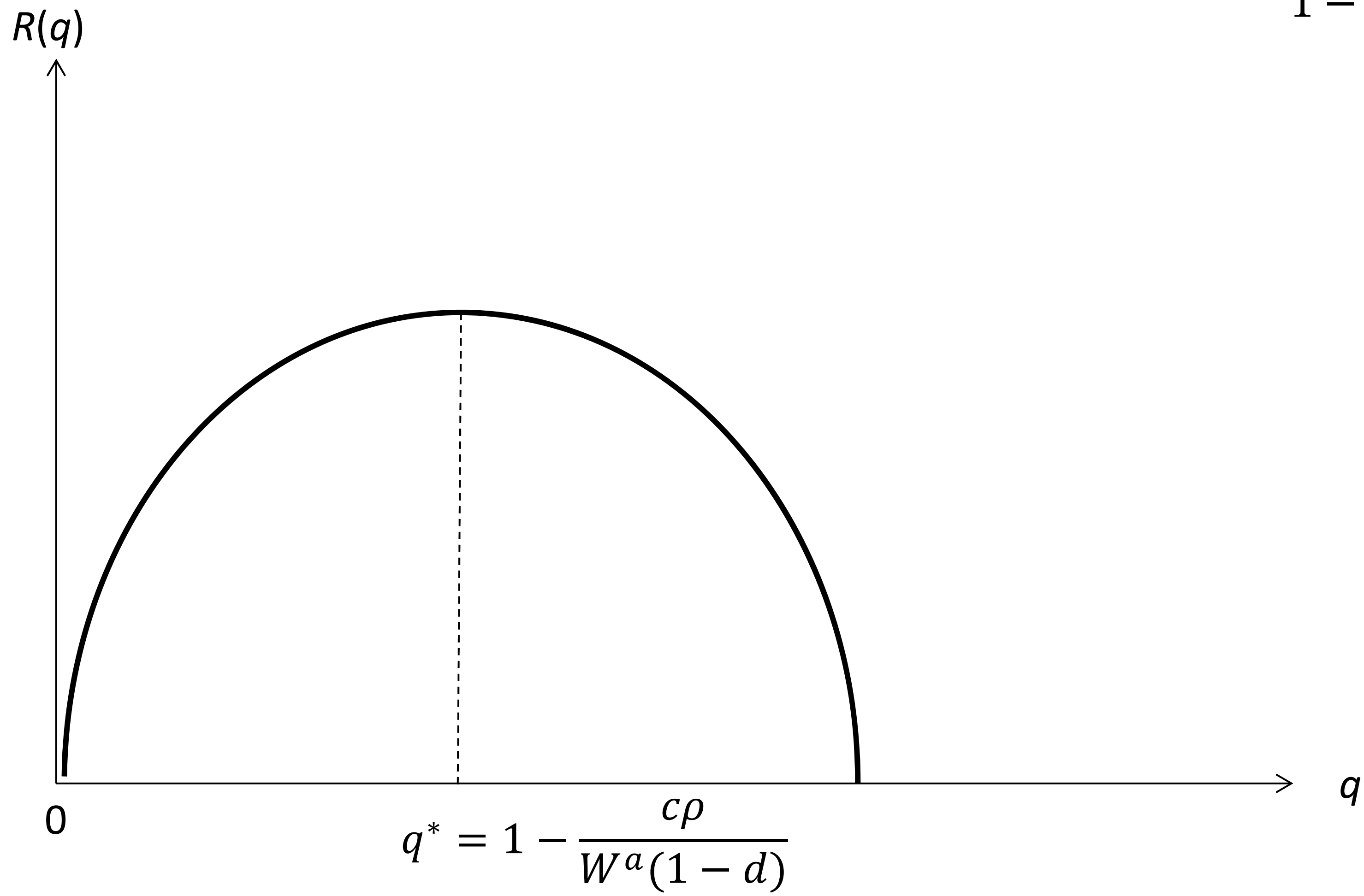

Figure D2: The choice of γ along the transition path when $W^a > c\rho/(1-d)$ and $q(0) < q^*$

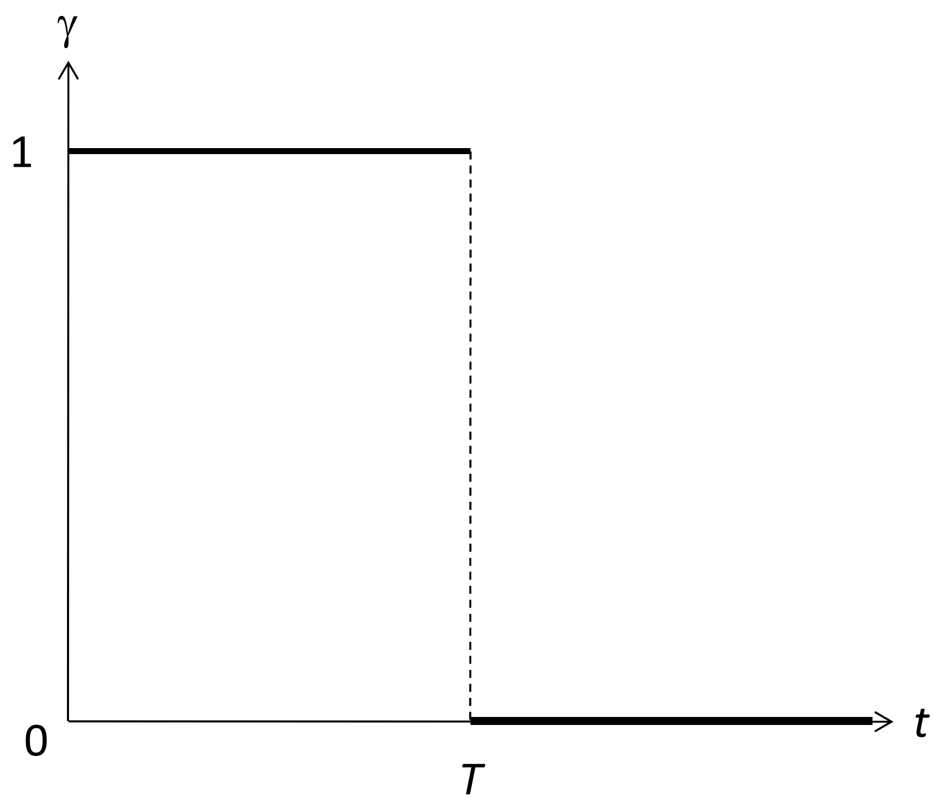

Figure D3: Case when $\frac{W^a}{c} > \frac{\rho - \Delta d}{1-d^b}$

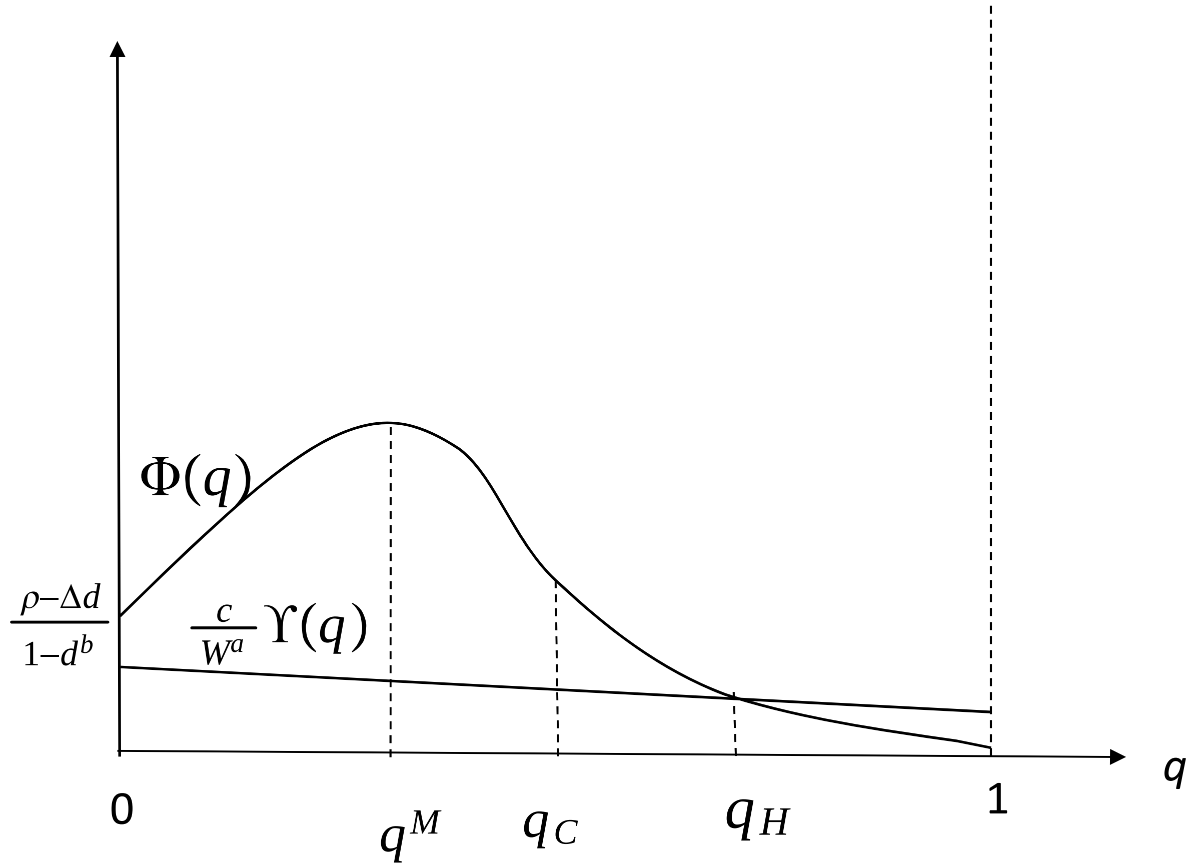

Figure D4: Case when $\frac{W^a}{c} \leq \frac{\rho - \Delta d}{1 - d^b}$

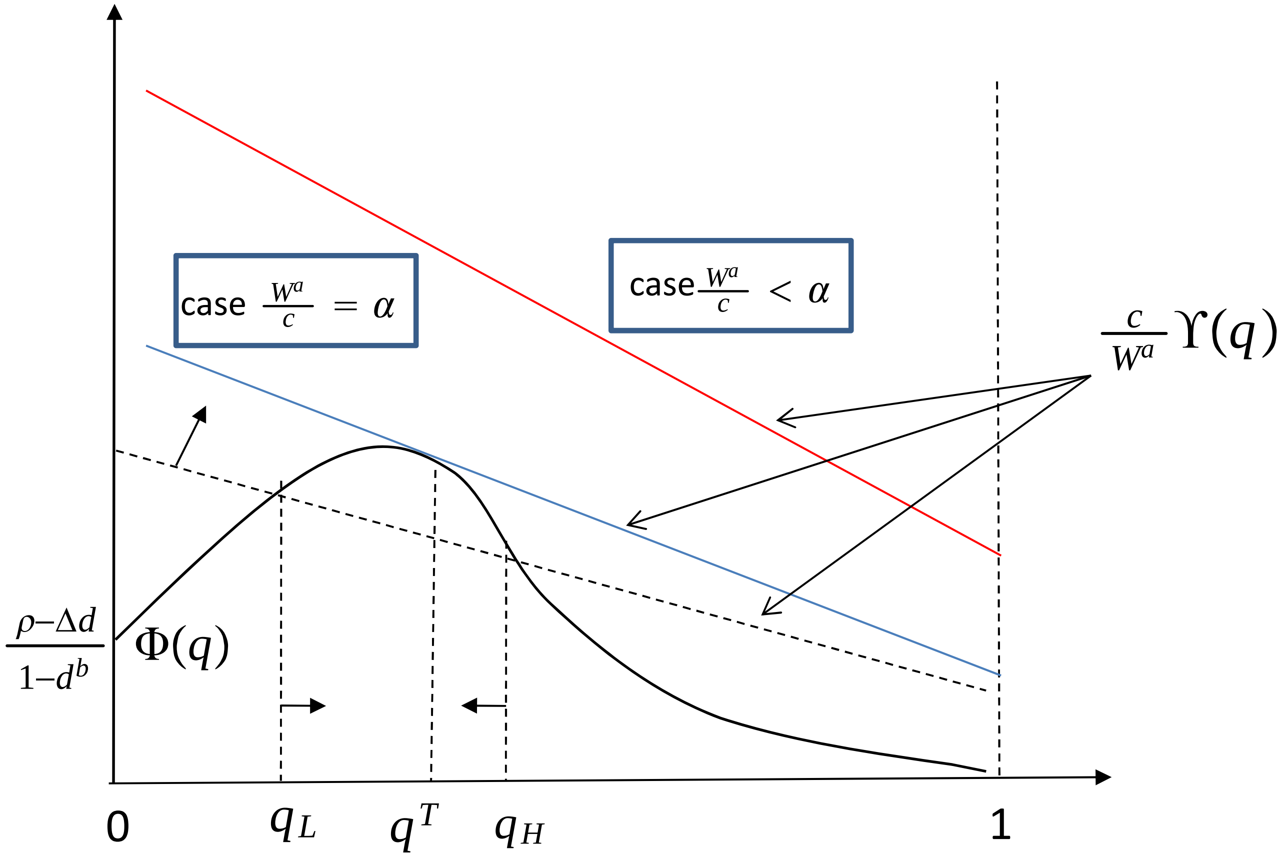

Figure D5 : Dynamics with leader competition