

# Community Leaders and the Preservation of Cultural Traits

Anja Prummer and Jan-Peter Siedlarek

Working Paper No. 806

December 2016

ISSN 1473-0278

## School of Economics and Finance



# COMMUNITY LEADERS AND THE PRESERVATION OF CULTURAL TRAITS\*

Anja Prummer<sup>†</sup>      Jan-Peter Siedlarek<sup>‡</sup>

November 23, 2016

## Abstract

We explain persistent differences in cultural traits of immigrant groups with the presence of community leaders. 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. With a leader cultural integration remains incomplete. The leader chooses more distinctive cultural traits in high productivity environments and if the community is more connected. Lump sum transfers to immigrants can hinder cultural integration. These findings are in line with integration patterns of various ethnic and religious groups.

**Keywords:** Cultural Transmission, Leadership, Immigrants, Labor Market Outcomes, Social Influence, Networks

**JEL Classification:** J15, Z10, D02

---

\*We are grateful to the Associate Editor and two anonymous referees for many useful suggestions on this paper and thank Árpád Ábrahám, Jérôme Adda, Alberto Bisin, Francis Bloch, Matt Elliott, Raquel Fernández, Sanjeev Goyal, Timo Hiller, Sriya Iyer, Rachel Kranton, Hamish Low, Massimo Morelli, Kaivan Munshi, Francesco Nava, Debraj Ray, Alireza Tahbaz-Salehi, Fernando Vega-Redondo and Yves Zenou for helpful discussions and advice. We also would like to thank seminar participants at Oxford University, University of Cambridge, SITE Stockholm School of Economics, NYU, INET-New York, CTN 2014, ASREC 2014, SAET 2014, EEA 2014 and the Workshop on Cultural Transmission and Evolution 2015. Siedlarek acknowledges financial support through DFG grant SFB-TR15. All remaining errors are ours.

<sup>†</sup>Corresponding author. School of Economics and Finance, Queen Mary University of London, Mile End Road, London, E1 4NS, UK, Telephone +44 (0)20 7882 5555, Email: a.prummer@qmul.ac.uk

<sup>‡</sup>Federal Reserve Bank of Cleveland, 1455 E 6th St, Cleveland, OH 44114, USA, Email: jan-peter.siedlarek@clev.frb.org. *Disclaimer:* The views stated herein are those of the authors and are not necessarily those of the Federal Reserve Bank of Cleveland or of the Board of Governors of the Federal Reserve System.

# 1 INTRODUCTION

It has been established since the 1960s that cultural integration is not a natural, inevitable process but that it regularly fails along religious lines ([Herberg \(1983\)](#), [Mayer \(1979\)](#)) as well as ethnic dimensions ([Glazer and Moynihan \(1963\)](#)). Nevertheless, cultural integration seems to be an important goal in many countries, such as the Netherlands, the US, France and Germany ([Brubaker \(2001\)](#), [Schalk-Soekar et al. \(2004\)](#)) who all pursue assimilationist policies. But how do immigrants adapt? When does cultural integration fail? Why does it often fail along *religious* and *ethnic* lines?

We address these questions and show how cultural and economic differences between immigrant communities and their host country persist in the long run, despite exposure to the culture of the host country and economic gains from assimilation.<sup>1</sup> We offer a novel explanation for this phenomenon, namely, community leaders who influence the identity of their community and thus act as identity entrepreneurs. Our notion of identity follows ([Akerlof and Kranton \(2000, 2010\)](#)) in that it denotes a sense of self. In the immigration context a higher identity corresponds to stronger identification with the culture of origin. This in turn implies that lower identity indicates greater assimilation. We show in a dynamic model of assimilation how leaders who benefit from a community that identifies with its origin culture prevent complete integration of immigrants.

Community leaders are common in immigrant communities. Despite this, the notion of a leader who creates a cultural identity is new to economics. We therefore shed some light on the leaders we consider, on who they are and in particular on their preferences. We show that they care both about the identity and the economic well-being of their community. Last, we present evidence that leaders are in fact able to influence the identity of their community.

We first turn to immigrant churches as they frequently fill the role of community leaders. As an illustrative example, we argue that in Turkish communities in Germany the leadership role falls to Imams and in particular the “DITIB”, an institution of the Turkish government. The DITIB employs the majority of Imams ([Yasar \(2012\)](#)) and can thus be seen as the ultimate leader. This institution follows a Turkish tradition of influencing cultural traits through Imams. Their impact was first recognized by Atatürk who harnessed them for his own political goals. He initially used Imams to mobilize the Turkish people in the

---

<sup>1</sup>As an example, gains from assimilation are higher earnings if an immigrant conforms to the norms prevalent in the host country, which in turn enables him to be more successful in the native labor market, see [Kantarevic \(2004\)](#).

Turkish War of Independence and later to shape the beliefs of the Turkish people, in particular emphasizing that a state-conform and moderate Islam was taught, often against the wishes of the Imams themselves. Atatürk achieved this by creating an organization that employed Imams directly and that ultimately led to the change he had aimed for (Ceylan (2010)).

Imams influenced the norms and values of their communities in historic Turkey, and they continue to do so in their communities in modern Germany: through their sermons and teachings they affect the assimilation of Turkish immigrants. According to Ceylan (2010, p.17), the political and religious orientation and the attitude of Imams towards the German government *decide* whether Muslims will be integrated in German society.<sup>2</sup>

Thus, it seems clear that Imams seek to spread their norms and values and therefore gain from a community that follows their teachings. At the same time, they also value monetary contributions from their community (Ceylan (2010, p.61)). Any religious movement seeks donations and encourages their members to contribute. It is well established that wealthier individuals donate more to their congregation, for an overview see Bekkers and Wiepking (2010). Thus, the Imams do not only care about the identity, but also the financial situation of their community.

Not only Imams (or for that matter Rabbis or Priests) can be considered to be leaders; foreign language media fill this role, too. For example, in Hispanic communities in the US, Spanish-language media often determine communication in the community and thus have an influence on the attitude towards integration (Portes and Sensenbrenner (1993)).<sup>3</sup> Like religious leaders, who are interested in the observance of religious cultural norms, foreign-language media have an interest in preserving language skills and the identity that comes with it and might even have incentives to discourage language acquisition as increased fluency in the host country language leads consumers to switch to host country media outlets (see Subervi-Velez (1986) for the Spanish-language media). Foreign-language media are businesses that are interested in maximizing their profits by having a large au-

---

<sup>2</sup>To be more precise about what features of norms and values Imams influence we consider the impact of being part of a Muslim community on the attitude to female labor force participation. Generally, those who identify with being Muslim, have a more traditional attitude towards female labor force participation, which in turn affects actual female labor force participation. Fortin (2005) shows that perceptions of women as homemakers are closely associated with women's labor market outcomes. Views that see men as the main breadwinners and women as homemakers are strongly influenced by religious ideology (Algan and Cahuc (2006); Guiso et al. (2003)). For example, Imams in Oslo emphasize that Islam does not forbid women to work in the public sphere, but that if women were to take outside jobs it should be in education or medical care (Predelli (2004)).

<sup>3</sup>Chinese-language media have the same role in Chinese communities (Zhou and Cai (2002)).

dience. They can achieve this by discouraging cultural integration, which prevents their consumers switching to the media of the host country. Additionally, they benefit from a wealthier community that can afford, for example, the newspapers produced or that is a more attractive target to advertisers.

What these examples highlight is that leaders benefit from their community and therefore have an interest in preserving a distinct group identity. They maintain such an identity by setting distinct norms and values that then influence the community's group members. Both religious leaders and foreign-language media care about the group members' norms as well as the economic well-being of their community.

We bring out the importance of community leaders in preventing cultural integration by first considering a group of immigrants where no leader is present. In this case the community will fully assimilate to the host country. We show this in a model of assimilation in discrete time. In each period, immigrants first decide how much to invest in host country specific skills (e.g. language skills and familiarity with norms and customs) which increase earnings as they improve the chances of immigrants in the host country labor market, taking their identity as given. A higher level of identity corresponding to a greater affinity with the country of origin makes it more costly to invest in skills and thus individuals with stronger identities invest less in skills.

In a second step, an immigrant's identity changes due to the skill investment.<sup>4</sup> Our notion of identity resembles the concept of a cultural trait. Therefore, we model the changes in identity, or rather in cultural traits, following the literature on continuous culture trait formation, a literature that goes back to [Cavalli-Sforza and Feldman \(1973\)](#).<sup>5</sup> In our setting, continuous traits are transmitted through a network.<sup>6</sup> This implies that an individual's identity is the weighted average of the host society's culture, his own past identity as well as the identity of the rest of the community, his network.

Departing from existing work on continuous cultural trait formation, skill investment affects the relative weight an individual puts on these sources and thus the weights change *in each period*. Higher skill levels imply a greater exposure to the host society which is reflected in the group member assigning a higher weight to the host society's culture. As

---

<sup>4</sup>The identity dynamics provide a departure from [Akerlof and Kranton \(2000, 2010\)](#), where identity remains fixed.

<sup>5</sup>Another key paper dealing with cultural trait formation is [Bisin and Verdier \(2000\)](#) who restrict attention to discrete cultural traits.

<sup>6</sup>Earlier work on continuous cultural trait formation in networks are [Büchel et al. \(2014\)](#) and [Panebianco \(2014\)](#).

a result, an immigrant with high skill levels will tend to identify less with the norms and values of his origin and more with those of the host country. In the subsequent period, this decrease in identity makes additional investment in skills less costly. An immigrant's assimilation process thus has a positive feedback loop: investment in skills reduces identity, which in turn, decreases the cost of investment, leading to higher investment in skills etc. As we show, in the long run, the immigrant group will be fully assimilated with the host country.

This result changes drastically when a community leader with a fixed identity is present. Then, assimilation will remain incomplete. In the presence of a leader, an immigrant's new identity is now also shaped by the identity set by the leader, in addition to the influence of the host society and that of the group members. Then, even though the group member invests in skill, his identity cannot converge to that of the host society as long as the leader has some influence on him. In particular, the long run identity of the group member lies strictly between that of the host society and that of the leader.

But which identity does the leader set? We characterize the economic and social environments in which the leader chooses full assimilation, intermediate assimilation or extremism. For the case of intermediate cultural integration, we also ask how different policy measures can affect the level of cultural distinction. Our results reverse what we find without the presence of a leader.

We first analyze the effect of an increase in the productivity in the environment. Greater productivity means that a given level of skill investment leads to greater economic returns holding all else equal. We find that if assimilation is intermediate, then being in a higher productivity environment induces the leader to set a higher level of cultural distinction. Higher productivity makes it easier to attain income and changes the trade-off between earnings and identity. This results in the leader countering the increase in productivity by setting a higher identity. Ultimately, in a high productivity environment, group members will have a higher identity, but will also be wealthier. Therefore, in this environment, immigrants have a higher welfare, with welfare depending on identity as well as earnings.

We can reinterpret this result as an insight on the level of discrimination in a society: a policy that reduces discrimination allows immigrants to access greater economic returns for a given level of skills investment and thus is equivalent to higher productivity for the immigrant groups. According to our model, such a decrease in discrimination will increase the community's welfare but will not lead to more cultural integration; in fact, it will reduce

it.

Further, we consider the effect of an economic support policy in the form of a lump sum transfer to immigrants on their assimilation outcomes. We find that in the presence of a leader a transfer can lead to less assimilation and less investment in host country specific skills. This highlights that economic wellbeing will not automatically lead to more cultural integration, a finding that is in line with [Krueger \(2008\)](#).

Last, we study the effect of community structure on assimilation. Group members influence each other and the extent to which they do so, that is, how much weight they assign to each other, also impacts the level of cultural distinction chosen by the leader. Specifically, the leader sets a higher level of identity when group members have a greater influence on each other. We interpret these type of groups as communities where individuals have strong ties, where individuals influence each other strongly. We offer thus a novel explanation of why groups with stronger ties assimilate less: namely, not only due to the ties per se, but also because they result in incentives for a leader to set a more distinct identity.

Our model can help to understand the assimilation patterns of different groups throughout history. We consider first the case of Jewish communities that assimilated fully in Germany after they were granted full citizen rights, but chose more cultural distinction in Hungary. This development can be attributed to Rabbis, who as the community leaders, faced different economic environments in these two countries. In particular, Germany was already industrialized at the time, whereas Hungary was still mainly agrarian. This led them to choose different paths for their respective communities as assimilation in Germany was more attractive as it increased earnings dramatically, whereas in a agrarian Hungary earnings would only have increased moderately.<sup>7</sup>

Second, we look at Muslim assimilation, which has been argued to be different from that of other ethnic and religious groups.<sup>8</sup> We show that possible rationales for such differences might be the greater influence of Imams relative to other religious leaders, a higher interest in cultural norms relative to economic success and strong bonds amongst family and community members.

**Related Literature** Our paper contributes to the literature that aims to understand different assimilation patterns of immigrant groups. Common explanations in the literature are par-

---

<sup>7</sup>For an overview of the Jewish assimilation, see [Carvalho and Koyama \(2011\)](#).

<sup>8</sup>For an overview, see [Bisin et al. \(2008\)](#), [Constant et al. \(2006\)](#), and [Haug \(2008\)](#).

ents' preferences for cultural traits (Bisin and Verdier (2000)),<sup>9</sup> ethnic and cultural distance to the host country (Alba and Nee (1997); Bisin et al. (2008)), previous educational background (Borjas (1985)) and discrimination against immigrants (Alba and Nee (1997)). We see our approach highlighting the role of group leaders as complementary to these explanations.

Methodologically, our paper is most closely related to the transmission of continuous cultural traits and builds on Cavalli-Sforza and Feldman (1973). Recent papers based on this approach include Büchel et al. (2014), Panebianco (2014) and Vaughan (2013). In these papers, the transmission of cultural traits occurs through parental choice, that is parents have a preference for their children to have the same cultural trait and therefore socialize them accordingly. In our setting, agents start out with given cultural traits and they choose to culturally integrate into the host society due to economic incentives. Agents shed their cultural traits to some extent as this increases their earnings in the host country labor market. Another key feature of all these papers is that with continuous cultural traits there is full assimilation in the long run unless there are persistent ties to the home country or subgroups are closed. In a different setting, Kuran and Sandholm (2008) also obtain full assimilation. Departing from this convergence result by taking into account the presence of a leader helps explain non-assimilation and differences in the assimilation processes of different immigrant groups (Bisin and Verdier (2010)).

The technical contribution of our paper derives from the feedback loop between skill investment and identity formation in our model. This interaction implies that the updating matrix describing the evolution of identity from one period to the next is time varying. We show that an intuitive sufficient condition on the relationship between the identities and changes in the updating matrix ensures convergence to a unique steady state. Our approach allows for all patterns of social interaction and does not require assumptions on the properties of the time-varying matrices, as has been the approach in the literature thus far.<sup>10</sup> Büchel et al. (2014) prove that their model converges with a time-varying updating matrix under certain restrictions on the updating matrix, which we do not need to impose. Panebianco (2014) similarly offers a convergence result for the case of time-varying updating matrix, which imposes a restriction that is not applicable to our settings. We discuss

---

<sup>9</sup>Fernández et al. (2004) and Fernández and Fogli (2006) also show that cultural traits are shaped by parents.

<sup>10</sup>DeMarzo et al. (2003) allow for some variation in the weight agents assign to their own identity, but ultimately they show that their specific setup has the same long run behavior as a model with a fixed updating matrix.

these technical aspects in more detail below Proposition 3.

Our technical contribution is thus two-fold: First, we deal with a case that is not covered by existing literature and present a complementary convergence result that does not use the assumptions made before. Instead, we exploit a different condition restricting the speed of change of the updating matrix as well as the specific social configuration with the leader and host society in our model. Second, our main convergence proof (Proposition 3) follows a novel approach making use of different tools to establish that the updating process is a contraction. This may turn out to be helpful to subsequent research in this field.

Furthermore, our paper is the first in the cultural transmission literature that studies the strategic incentives of a leader in the choice of identity.<sup>11</sup> The importance of community leaders has also been documented in [Carvalho and Koyama \(2011\)](#). The key difference to their work is that our model allows for different types of leaders, whereas they only look at religious leaders, namely at Rabbis in Jewish communities. We generalize their payoff function and we explicitly model the assimilation process and take group structures into account. Further, [Hauk and Mueller \(2015\)](#) discuss the role of cultural leaders in the clash of cultures. In their setting the leader is not crucial for the clash, contrary to our finding. Additionally, our model adds considerations regarding group structure and the impact of policies on the strategy of the leader.<sup>12</sup> Within the area of cultural transmission, subsequent work by [Verdier and Zenou \(2015\)](#) introduces a second strategic cultural leader and studies the resulting contest between leaders. Their analysis differs from ours by adopting a [Bisin and Verdier \(2000\)](#) framework with discrete cultural traits, whereas in our model cultural traits are continuous. In addition, their framework does not allow for heterogeneity in the cultural integration of individuals, skill investment or an explicit network structure. Despite these differences, [Verdier and Zenou \(2015\)](#) confirm that the presence of leaders can prevent full integration, underlining the robustness of our finding.

The remainder of this paper is structured as follows. We present the model of the assimilation process in Section 2 and our analysis in Section 3. We first analyze assimilation without the presence of a leader as a benchmark in Section 3.1. Section 3.2 then shows how a leader can prevent integration. In Section 4, we allow the leader to be strategic and study implications for assimilation outcomes. Comparative statics and policy implications are discussed

---

<sup>11</sup>Other work on opinion dynamics includes [Acemoglu et al. \(2010\)](#), [Acemoglu et al. \(2013\)](#), [Böchel et al. \(2015\)](#), and [Lorenz \(2006\)](#).

<sup>12</sup>Empirically, [Munshi and Wilson \(2008\)](#) have also emphasized the importance of churches in the preservation of cultural traits.

in Section 5. Section 6 considers a variant of our model that allows agents to sever ties to their community and shows that our findings carry over to this setting. Section 7 connects our findings to a number of applications. Section 8 concludes. All proofs are collected in the Appendix.

## 2 A MODEL OF COMMUNITY ASSIMILATION

In our model there are  $n$  group members which are influenced by the host society  $S$  and the group's leader  $L$ .<sup>13</sup> The host society does not take any action in our model; only the group members and the group leader are active players.

The model is dynamic and set in discrete time. Group members are characterized at any point  $t$  by their level of skills  $H_i^t \in \mathbb{R}_0^+$  and their identity  $p_i^t \in [0, p^{\max}]$ . The identity measure captures the attachment an individual has with his group, with higher  $p_i^t$  indicating higher levels of group attachment and greater differentiation from the host society. We bound the space of permissible identities and assume that there is an upper level  $p^{\max}$ .<sup>14</sup> We fix  $p_S^t = p_S = 0$  for the host society for all time periods.<sup>15</sup> In what follows we focus on long run outcomes in which the leader sets the optimal identity,  $p_L \in [0, p^{\max}]$ , for the steady state.<sup>16</sup>

A group member  $i$  is fully identified with the host country if  $p_i^t = 0$ , whereas  $p_i^t = p_L$  indicates full identification with the group leader. Group members invest in skills  $H_i$  and adapt their identity over time taking as given the leader's policy. We first consider the assimilation process of the group members for a given group leader policy  $p_L$  and then turn to the resulting incentives for the leader.

<sup>13</sup>Note that we focus on the case in which there is a single leader in the community acting as a cultural monopolist. While we think this is a natural setting to study and relevant for many applications, the study of a model of cultural transmission with competing leaders appears to be a fruitful area for future research.

<sup>14</sup>The upper bound  $p^{\max}$  can be seen as a restriction imposed by the host country. This upper bound can be normalized but this does not present a significant simplification and thus we leave it free.

<sup>15</sup>We assume that the host society is not influenced by the immigrant group. As such our model focuses on one side of the assimilation process. We consider this an innocuous assumption as in many settings immigrants make up only a small percentage of the population. Immigrants might have a local effect, but will in most cases not change a country's culture. Furthermore, by keeping the overall identity of the host country fixed we are able to specifically study the identity dynamics of interest, namely of the immigrant group.

<sup>16</sup>One reason for this could be that the leader is infinitely patient and therefore only cares about long-run outcomes. Note that this does not imply that it is never optimal for the leader to adjust identity. If there is a change in the environment which leads to a different steady state, then the leader may find it beneficial to adjust his position for this new environment. More precisely, we focus on the optimal steady state  $p_L$  for a given environment as the long run outcome and we do not investigate in detail the leader identity path approaching the steady state. A recent follow-up paper that explores the dynamics of cultural leaders positioning over time away from the steady state is [Verdier and Zenou \(2016\)](#).

## 2.1 Group Member Assimilation

In every period  $t$ , each group member goes first through a process of identity adaptation and in a second step invests in skills. We discuss each of these processes in turn.

**Identity Dynamics** Group members do not actively choose their identity; instead it is adjusted passively with reference to the social environment. This is in line with the standard assumption in [Akerlof and Kranton \(2010\)](#) where identity is not a choice variable, but rather a preference or type.<sup>17</sup> A group member's identity is composed as a weighted average of the values and norms of (i) the host society and (ii) the leader as well as (iii) the past identity of the group members themselves (including member  $i$ 's past identity). The weights on these three sources of influence are determined as follows:

(i) *Host Society* Each group member is influenced by the host society,  $S$ . We denote the share of influence given to that source by  $g(H_i^{t-1}) \in (0, 1)$  which is a strictly increasing function of the previous period's investment in skills,  $H_i^{t-1}$ . This captures the fact that with greater levels of investment, the group member is more exposed to the influence of the host society, for example through new social connections or greater exposure to the host country media. It also implies that the weight on the host society changes over time.

(ii) *Leader* Of the residual, the leader captures a share  $\lambda_i \in (0, 1)$ . The overall weight on the leader is then given by  $\lambda_i(1 - g(H_i^{t-1}))$ . The parameter  $\lambda_i$  is an indicator measuring the influence of the leader on group member  $i$ , with higher  $\lambda_i$  indicating a more influential leader.

(iii) *Group Members* The weight group member  $i$  assigns to group member  $j$  is denoted by  $\gamma_{ij}(1 - \lambda_i)(1 - g(H_i^{t-1}))$ , where  $\gamma_{ij} \in [0, 1]$  presents the relative within-group weight  $i$  puts on  $j$  such that  $\sum_{j=1}^n \gamma_{ij} = 1 \forall i$ . The weights between group members represent the strength of their social connections and they vary over time, depending on the level of skill investment.

An overview of the process is provided in [Figure 1](#). The identity adaptation process for group member  $i$  from  $p_i^{t-1}$  to  $p_i^t$  can then be summarized as follows,

$$p_i^t = [1 - g(H_i^{t-1})] \cdot \left\{ \lambda_i p_L + (1 - \lambda_i) \sum_{j=1}^n \gamma_{ij} p_j^{t-1} \right\}$$

where we have used the fact that  $p_S = 0$  so that the influence term of the host society

<sup>17</sup>[Bisin and Verdier \(2001\)](#) argue that cultural traits are already formed very early in life (kindergarten), which makes it plausible to take identity as given when skills are chosen.

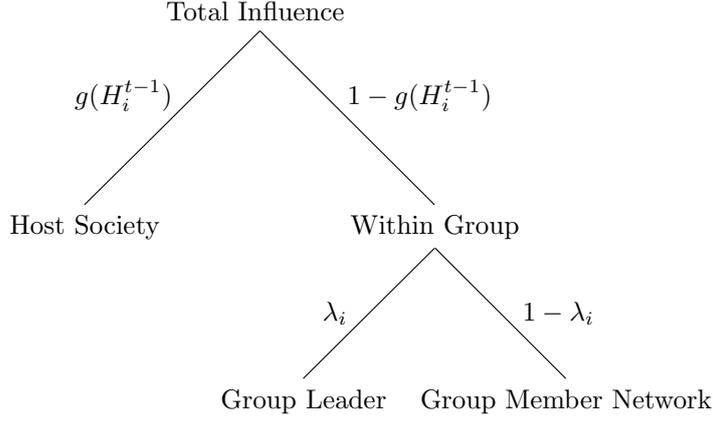


Figure 1: Influence Weights in Period  $t$

vanishes. Note that the normalization of the host society's identity to zero is without loss of generality. We refer to the *symmetric* case as the setting with two group members who each assign the same weight  $\gamma \in [0, 1]$  to themselves and  $(1 - \gamma)$  on the other group member. In this case, we refer to  $\gamma$  as group division as it is a measure of how divided the group is. By the same logic  $(1 - \gamma)$  is then a measure of group cohesion. Setting  $\gamma = 1$  reflects the special case of the group member being isolated from the influence of other group members and therefore maximal group division.

**Skills Investment** In a second step, group members invest in skills  $H_i^t \in \mathbb{R}_0^+$ . This measures a group member's effort to learn the language, to understand the norms and cultures of the host society, which leads to higher income in the host country labor market.<sup>18</sup>

Each period the group members simultaneously select how much to invest in skills. The payoff from skills  $H_i^t$  is given by  $\alpha_i f(H_i^t)$ . The function  $f(H_i^t)$  is strictly increasing and concave in  $H_i^t$ . We also assume that it is continuous and three times continuously differentiable. Additionally, payoffs from investment depend on the parameter  $\alpha_i \in (0, 1)$ , which reflects productivity.

Investing in skills is costly with the cost depending on the identity of the group member. For ease of exposition we let this cost be linear in the investment level with marginal costs  $c(p_i^t)$ . We assume that marginal costs are strictly increasing in  $p_i^t$ ,  $c'(p_i^t) > 0$ , capturing the notion that immigrants who are more deeply rooted in their culture and whose norms and values are more different from the host society face higher costs understanding and accessing their new environment.

<sup>18</sup>Skills should not be interpreted as schooling or human capital that is useful both in the immigrant and the native labor market, but rather as skills that only give an advantage in the native labor market compared to the immigrant one.

Finally, we assume that the marginal net return to additional investment at  $H_i^t = 0$  is positive for all agents and levels of group identification, i.e. for all  $i$  and  $p_i^t$ ,  $\alpha_i f'(0) > c(p_i^t)$ . This implies that all group members will choose a positive level of investment in skills for any given level of group identity.

In summary, in each period  $t$ , group members have an identity  $p_i^t$  and a stock of skills,  $H_i^t$ , which determines their earnings or wealth  $\alpha_i f(H_i^t)$ . We now turn attention to the question how these two variables determine the leader's payoff.

## 2.2 The Leader's Decision Problem

The leader selects a level of identity  $p_L \in [0, p^{\max}]$  to maximize his payoff. We specify the per period payoff function that depends on the identity and the wealth or earnings,  $\alpha_i f(H_i^t)$ , of the group members as follows:

$$\sum_{i=1}^n \pi(p_i^t, \alpha_i f(H_i^t))$$

where  $\pi_1(p_i^t, \alpha_i f(H_i^t)) > 0$ ,  $\pi_2(p_i^t, \alpha_i f(H_i^t)) > 0$ , that is, the leader's payoffs are increasing in identity as well as earnings.<sup>19</sup> We further assume that

$$\pi_{11}(p_i^t, \alpha_i f(H_i^t)) \leq 0, \quad \pi_{22}(p_i^t, \alpha_i f(H_i^t)) \leq 0, \quad \pi_{12}(p_i^t, \alpha_i f(H_i^t)) \geq 0.$$

This implies that the leader payoffs are weakly concave in both identity and earnings. Additionally, earnings and identity are (weak) strategic complements in the leader's utility function.

Motivated by the examples provided in the introduction, we consider in our analysis two special cases, namely that of a *religious* leader (*RL*), who cares about the identity per se and a *business* leader (*BL*), who cares only about identity as it enhances the profits he obtains from the community. One way to specify these payoff functions is given in equations (1) and (2):<sup>20</sup>

$$\pi_{RL}(p_i^t, H_i^t) = \beta p_i^t + (1 - \beta) \alpha_i f(H_i^t), \quad (1)$$

$$\pi_{BL}(p_i^t, H_i^t) = (p_i^t)^\beta (\alpha_i f(H_i^t))^{1-\beta}, \quad (2)$$

<sup>19</sup>As mentioned previously, we focus on steady state outcomes in our analysis and therefore the identities in and earnings can be considered as starting in steady state.

<sup>20</sup>We follow [Carvalho and Koyama \(2011\)](#) in their specification of the payoff function of the religious leader.

Equations (1) and (2) can be seen as two special cases of a CES utility function,  $(\beta(p_i^t)^\sigma + (1 - \beta)(\alpha_i f(H_i^t))^\sigma)^{\frac{1}{\sigma}}$ , with  $\sigma = 1$  and  $\sigma \rightarrow 0$ . The parameter  $\sigma$  governs the extent to which the leader sees identity and earnings as strategic complements.

**Discussion: Leader Payoff** We assume that the leader prefers a community that is as culturally distinct from the host society as possible, but at the same time as wealthy as possible. This introduces the key trade off of the paper. If the leader only cares about identity, then it is straightforward to show that he always sets the highest identity possible,  $p_L = p^{\max}$ . On the other hand, if the leader only cares about the wealth of the community, then he would set his identity to zero, in order to foster complete cultural integration. Therefore we focus our attention to the most interesting case of the leader caring about both identity and earnings. To motivate this further, we discuss a few of the numerous examples that document that community leaders indeed care about both identity and earnings.

*Religious Organizations* Religion clearly impacts the identity of its followers. This can be easily observed in daily life when considering the norms that different religious groups display. Peyes and headscarfs tend to identify the religion of the wearer and thus shape their identity. The effect of religion in creating specific identities is widely documented, see for example [Breton \(1964\)](#) and [Carvalho and Koyama \(2011\)](#). Furthermore, leaders care about the financial wellbeing of their community. This might be due to a need to maximize donations ([Ceylan \(2010\)](#); [Bekkers and Wiepking \(2010\)](#)) or from a sense of prestige or altruism. Every religious organization we are aware of asks for donations. A religious leader benefits from these donations as this allows for example to build a better church, synagogue, mosque or temple. We would further expect a religious leader to care not only about the spiritual well-being of the community members, but also about their material well-being. We believe that most religions do want their followers to be able to at least afford food, clothing as well as housing.

*Foreign Language Media* Different from religious institutions, media outlets are profit maximizing firms. Foreign language media can increase their profits by creating a specific identity that leads consumers to seek them out more. An example of this is given in [Portes and Sensenbrenner \(1993\)](#) regarding Latino groups in South Florida, who are described as completely controlled by three radio stations which determine the identity within this community. Due to the difference in identity between Hispanics and US Americans, Hispanics seek out media outlets that provide programs in Spanish, leading to higher profits for such

outlets (Subervi-Velez (1986)). The same patterns can also be found for Chinese language media (Zhou and Cai (2002)). Further, media outlets finance themselves mainly through advertising revenue and therefore need to have an audience that is worth being advertised to. If the community is poor, then it is doubtful that it is a valuable market for businesses. Therefore, media outlets also have an interest in their community being wealthy. An interesting example of this is described in Zhou and Cai (2002). They show that Chinese media outlets in the US provide advice to their audience on business opportunities and how to achieve financial success.

Based on these examples, we argue that community leaders prefer a community that is both identified with the norms and values propagated as well as wealthy. Additionally, there is evidence that individuals who are more identified with a community are also willing to contribute more financially (Bekkers and Wiepking (2010)). We therefore impose the assumption of identity and earnings being strategic complements.<sup>21</sup> To simplify our analysis further and in line with the examples provided, we consider a general CES utility function for the leader. This allows us to consider different types of leaders, that is religious versus business-oriented leaders, such as foreign-language media. Note however that none of our results depend on the specific functional forms we impose and the example of the religious and business leader only serve to illustrate these results.

As argued previously, it seems plausible that religious leaders are interested in the identity of their community per se. A natural way to model this is to allow for the religious leader to have a payoff function that does not have complementarities between identity and earnings. Different from the religious leader, a business does not care about the creation of an identity per se. The business cares about identity as it allows to realize higher earnings, but this is only possible if the follower has a sufficient income. It is natural to capture this consideration by assuming that identity and earnings are strategic complements.

### 3 ASSIMILATION DYNAMICS WITH AND WITHOUT LEADER

We first study as a benchmark the assimilation process of a community without leader. This serves to document the forces towards assimilation in our model and thus highlights why a leader is crucial to preserve the identity of an immigrant group in the long-run. We then turn to assimilation in the presence of a leader.

---

<sup>21</sup>We provide additional results in the Appendix for the case of strategic substitutes. However, it is not clear to us how one could motivate or interpret a case of strategic substitutes in this setting.

### 3.1 Benchmark Case: Identity Dynamics without Leader

We consider in turn the optimal skill investment and identity adaptation.

**Optimal Skills Investment** Each period group members select a level of skills investment based on the level of group identity they have in this period.<sup>22</sup> The optimal investment is given by the solution of the following maximization problem:

$$\max_{H_i^t \geq 0} \{ \alpha_i f(H_i^t) - c(p_i^t) H_i^t \}.$$

This set-up assumes that group members are myopic which we see as consistent with the interpretation of identity as a preference for a certain type of behavior. We study in the Appendix two alternative models with forward looking agents and show that this assumption does not affect our qualitative results. We solve the maximization problem using the first order condition, that is both necessary and sufficient and identifies a unique interior solution.

We write the optimal skill investment as a function  $H^*(p_i^t; \alpha_i)$ . The optimal investment level is decreasing in  $p_i^t$ . As group members identify more with their home group, their desired level of skill investment decreases, reflecting the greater costs of such assimilation efforts. Furthermore, agents with higher  $\alpha_i$  have higher investment levels for any given identity level  $p_i^t$ . As the returns to skills investment increase for immigrants due to higher productivity, they find it beneficial to invest more.

**Identity Dynamics** Based on their skill investment in period  $t$ , group members update their identity  $p_i^{t+1}$ . Recall that the weight given to the host society is given by  $g(H_i^t)$  with  $H_i^t$  chosen through function  $H^*(p_i^t; \alpha_i)$ . Taking this into account we define

$$\hat{g}(p_i^t; \alpha_i) \equiv g(H^*(p_i^t; \alpha_i)).$$

This function  $\hat{g}(p_i^t; \alpha_i)$  maps every identity level  $p_i^t$  into  $(0, 1)$  for a given  $\alpha_i$ . It is *decreasing* in  $p_i^t$  as  $H^*(p_i^t; \alpha_i)$  is a decreasing function in  $p_i^t$  and  $g(H_i^t)$  is increasing in  $H_i^t$ . Furthermore, for every  $p_i^t$  it is increasing in  $\alpha_i$  as a higher  $\alpha_i$  implies a higher  $H^*(p_i^t; \alpha_i)$ , which leads to a

---

<sup>22</sup>In our specification, we ignore that group members might derive utility from their identity. However, one can easily adjust the maximization problem and include a term that depends on identity. This will not change our maximization problem as long as this term enters additively.

higher weight on the host society.

Next period's identity,  $p_i^{t+1}$ , can then be written as a function of the previous period identity levels as follows:

$$p_i^{t+1} = [1 - \hat{g}(p_i^t; \alpha_i)] \left\{ \sum_{j=1}^n \gamma_{ij} p_j^t \right\} \quad (3)$$

We focus here on the long run assimilation outcome, by studying the steady state of the system. The steady state identity vector  $\bar{p}$  is characterized by constant identity levels for each group member that satisfy

$$\bar{p}_i = [1 - \hat{g}(\bar{p}_i; \alpha_i)] \left\{ \sum_{j=1}^n \gamma_{ij} \bar{p}_j \right\}. \quad (4)$$

The corresponding levels of steady state investment  $\bar{H}_i$  is equivalent to the optimal skill investment,  $H^*(\bar{p}_i; \alpha_i)$ . As investment is strictly positive for all identity levels, in the steady state group members put a strictly positive amount of weight on the host society. We now use this property to fully characterize the steady state identity vector for the benchmark case without leader.

**Proposition 1** (Steady State without Leader).

*In the steady state without a leader group members assimilate fully with long run identities converging to zero.*

Every group member has a positive level of investment in skills in each period, which leads to a strictly positive weight on the host society. A greater weight on the host society then leads step-by-step to a lower identification with the group. This might initially be countered by the network, as a group member might be connected to other group members that have a higher identity. Therefore, through the influence of peers, identity could initially increase. Nonetheless, in the long run, all immigrants will be fully assimilated. Note that our result is not driven by  $\hat{g}(0; \alpha_i)$  approaching one or the fact that all group members invest in skills but owed to the fact that gradually the identity of all group members decreases. In particular, the only assumption necessary for this result is that in every period  $\hat{g}(p_i^t; \alpha_i)$  lies between zero and one and is independent of the functional form of  $\hat{g}(p_i^t; \alpha_i)$ .

A consequence of this is that our result carries over to a setting where some group members never invest in skills. This can be the case if the group member's payoff function is

given by

$$\max \{ \alpha_i f(H_i^t) - c(p_i^t) H_i^t, o \}, \quad (5)$$

where  $o$  denotes an outside option.<sup>23</sup> To motivate this, consider an immigrant who can choose to either invest in skills, which improve his chances in the host country labor market or to pursue a career within his immigrant community. The job within his community does not require any skill investment and pays an amount of  $o$ . For low productivity types it might then never be optimal to invest in host country specific skills. This does not have an effect on the long run assimilation outcome, as each group member is still influenced by the host society as  $g(H) \in (0, 1)$  for any level of skill investment.

We turn to the convergence process next. We discuss the speed at which group members assimilate if they assign equal weights to other group members as a function of their initial identity and their productivity. Our result is summarized in the following proposition.

**Proposition 2.** *Consider a symmetric setting with  $k$  group members that assign weight  $\gamma \in [\frac{1}{k}, 1]$  to their own identity and equal weight  $(1 - \gamma) / (k - 1)$  to each of the  $k - 1$  other members.*

1. *If all group members have the same initial integration level, then group members of high ability integrate faster than those of lower ability, in the sense that at any point of time their identity is lower than that of group members of lower ability.*
2. *If group members have the same ability, then group members with lower initial identity integrate faster than those with higher initial identity, in the sense that at any point of time their identity is lower than that of group members with higher initial identity.*

If we consider the case of isolated group members ( $\gamma = 1$ ), then our result is fairly straightforward. Group members whose initial identity differs more from that of the host country face a slower assimilation process than those whose initial identity is more similar. Additionally, individuals who have a higher productivity assimilate quicker. For them, it is more profitable to invest in skills and therefore at each point in time, they invest more in skills. This speeds up their assimilation process. If group members assign positive weights to each other, this pattern does not necessarily hold. For example, consider an extreme case with a high and low ability group member that assign weight 1 to each other, and no weight on themselves. Then, even if initially the identities of both group members are identical, it

---

<sup>23</sup>This formulation is mathematically equivalent to a discrete production function  $f(H)$  which does not pay anything unless  $H$  is sufficiently high, that is a production function that only pays once immigrants have made a sufficiently high skill investment.

can be the case that the low ability group member has a greater convergence speed as it is more influenced by the high ability group member than the high ability agent himself. The assumption that  $\gamma$ , the weight on oneself, has to be sufficiently high as well as the equal weight assumption on all neighbors ensures that such a pattern cannot occur.

### 3.2 Identity Dynamics with Leader

We now introduce a leader with some fixed  $p_L$  and obtain the steady state identities of the group members. Based on these steady state identities, the leader then sets the optimal identity.<sup>24</sup>

The updating process of group member  $i$ 's identity is given by

$$p_i^{t+1} = [1 - \hat{g}(p_i^t; \alpha_i)] \left[ \lambda_i p_L + (1 - \lambda_i) \sum_{j=1}^n \gamma_{ij} p_j^t \right]. \quad (6)$$

To guarantee uniqueness of steady state as well as global convergence of the system, we impose an additional assumption on  $\hat{g}(p_i; \alpha_i)$ .

#### Assumption 1.

$$\left| \frac{\partial \hat{g}(p_i; \alpha_i)}{\partial p_i} \right| \frac{p_i}{\hat{g}(p_i; \alpha_i)} < \frac{p_i}{p^{\max}} \quad \forall p_i, \alpha_i$$

The elasticity of  $\hat{g}(p_i; \alpha_i)$ ,  $\left| \frac{\partial \hat{g}(p_i; \alpha_i)}{\partial p_i} \right| \frac{p_i}{\hat{g}(p_i; \alpha_i)}$  has to be smaller than the fraction  $p_i$  over  $p^{\max}$ . This assumption guarantees that the identity adjustment process is sufficiently smooth and small changes in identity today do not have too large of an impact tomorrow. Intuitively, it requires that the weight that group members assign to their group does not decline too rapidly in response to small decreases in their identity  $p_i$ . The assumption does not rely on the fact that group members choose skill investment optimally given their identity. All we need is a function that depends on identity which fulfills the property of Assumption 1. How we arrive at this function is per se irrelevant.

We then establish the following result concerning the long run outcome of the setting with a leader.

#### Proposition 3 (Steady State with Leader).

<sup>24</sup>Note that we could easily turn this into a one period problem, where we simply focus on the steady state identities on the group members without deriving them explicitly. We believe that such an approach would be rather ad hoc and therefore choose to show that such a steady state in fact exists and when it is unique.

*In the model with a leader in the influence network:*

1. *A steady state exists.*
2. *There is no full assimilation into the host society. In every steady state, a group member's identity is a strictly convex combination of the position of the host society and that of the leader.*
3. *If Assumption 1 holds the steady state is unique and the group member identities converge globally.*

Proposition 3 establishes that a steady state always exists. Further, Assumption 1 ensures that the steady state is unique and group member identities converge globally. Our convergence result here applies to a setting not studied elsewhere and thereby complements and extends other results in the literature.<sup>25</sup> Specifically, Büchel et al. (2014) require the updating matrix to be symmetric ultrametric to establish convergence in their setting. We do not need such restrictions and as they exclude any inessential communication classes their methodology cannot be extended to our setting.<sup>26</sup> Panebianco (2014) studies a model with multiple traits and vertical as well as oblique socialization. There are two main differences to our model. First, in our setting, we do not impose a condition requiring strictly positive diagonals in our updating matrix, while this property is key in Panebianco (2014)'s analysis. There an assumption on payoffs ensures that parents always put strictly positive effort into socialization (as shown in Proposition 1 of Panebianco (2014)). This property is then a key element of the proof of Proposition 2 (p. 603) in Panebianco (2014), ensuring acyclic updating matrices. We do not impose this restriction and in principle allow for zero diagonal entries for group members. Second, our model permits greater time variation than the main convergence result in Panebianco (2014) (Proposition 2), which permits time-variation only on the diagonal. Corollary 1 in Panebianco (2014) allows for more general time variation, but this result restricts attention to cases with one communication class. Our model with leader as analyzed in Proposition 3, however, inherently has at least two communication classes. Thus, the results from Panebianco (2014) do not translate to our setting.

In addition to dealing with a case not studied elsewhere, we also offer a novel approach

---

<sup>25</sup>The characterization of the steady state outcome as a convex combination of the identity of the host society and the leader mirrors DeMarzo et al. (2003, Theorem 10). However, they derive the result for the time-homogeneous case only.

<sup>26</sup>See Büchel et al. (2014) for a definition of inessential communication class. Intuitively, a communication class is essential if each member of the class is only connected to other members of the class, but not to agents outside the class. If a communication is not essential, it is inessential. See also Lorenz (2005) for further background on this approach.

to establish convergence and uniqueness of the long-run steady state in the class of cultural transmission models. As we show in detail in the Appendix, the elasticity condition Assumption 1 ensures that the updating process forms a contraction using a suitable metric and therefore converges to a unique steady state from any given starting point, avoiding multiplicity of long-run outcomes. If Assumption 1 is violated in the sense that the influence assigned to the host society is too responsive to changes in group member identity, we cannot exclude the presence of cycling and multiple steady states.

In terms of our application, the result shows that the presence of a leader, who is not susceptible to any influence, generally prevents full assimilation.<sup>27</sup> However, the group members will assimilate to some extent in *any* steady state.

If we consider the case where group members have a discrete payoff function as specified in equation (5), then our results remain unchanged. Assumption 1 is still satisfied, as now a change in  $p$  might either not trigger a change in the weight placed on the host society or the change is as described before, depending on whether there is positive skill investment or not.

**Network Structure and Integration** The long-run assimilation patterns of a community depend on the network structure within the group. In order to be able to characterize the interplay between the network and the steady state identities, we first require additional notation.  $\mathbf{I}$  is the  $n \times n$  identity matrix, and  $\mathbf{\Lambda}$  and  $\hat{\mathbf{G}}(p)$  are  $n \times n$  diagonal matrices with element  $ii$  given by  $\lambda_i$  and  $\hat{g}(p_i^t; \alpha_i)$ , respectively. The network structure is characterized by the  $n \times n$  matrix  $\mathbf{\Gamma}$ , which contains the weights group members assign to each other and themselves,  $\gamma_{ij}$ . Last,  $\mathbf{1}$  denotes a  $n$  column vector of ones.

**Proposition 4.** *For an arbitrary influence network  $\mathbf{\Gamma}$ , the steady state identity vector for group members is characterized by*

$$\bar{p} = \left\{ \mathbf{I} - (\mathbf{I} - \mathbf{\Lambda}) \left[ \mathbf{I} - \hat{\mathbf{G}}(\bar{p}) \right] \mathbf{\Gamma} \right\}^{-1} [\mathbf{I} - \hat{\mathbf{G}}(\bar{p})] \mathbf{\Lambda} \mathbf{1} p_L. \quad (7)$$

Equation (7) brings out the role of the structure of the network  $\mathbf{\Gamma}$  that captures the social relations amongst the members of the group. The steady state identities  $\bar{p}$  are pinned down by an equation that mirrors the weighted Bonacich centrality, a concept highlighting the

---

<sup>27</sup>If the leader were susceptible to influence by the group members, the leader would place a strictly positive weight on their identity. This is mathematically equivalent to the leader not being present at all, and we would end up with full integration into the host society in steady state as per Proposition 1.

importance of a node in a network.<sup>28</sup> In our setting, however, the network structure is less important for assimilation outcomes compared to individual productivity and the influence of the leader.

Note first that in the second half of the expression the leader's identity  $p_L$  is multiplied by  $[\mathbf{I} - \hat{\mathbf{G}}(\bar{p})]\mathbf{\Lambda}$ . Thus for each group member  $i$ , this term depends on the influence the leader has on that member,  $\lambda_i$ , as well as the weight the group member assigns to the community as a whole relative to the host society,  $[1 - \hat{g}(\bar{p}_i, \alpha_i)]$ . Crucial for this weight is the productivity  $\alpha_i$ , which leads higher ability group members to assign a lower weight to the leader. As this part of equation (7) depends only on a group member's individual characteristics and is independent of other group members, we define this half of the expression as the  $n \times 1$  vector of "idiosyncratic" identities.

The network structure only comes into play in the first part of equation (7). Note that this term is an  $n \times n$  matrix and can be rewritten and reinterpreted as an infinite sum:

$$\left\{ \mathbf{I} - (\mathbf{I} - \mathbf{\Lambda}) \left[ \mathbf{I} - \hat{\mathbf{G}}(\bar{p}) \right] \mathbf{\Gamma} \right\}^{-1} = \sum_{k=0}^{\infty} \left\{ (\mathbf{I} - \mathbf{\Lambda}) \left[ \mathbf{I} - \hat{\mathbf{G}}(\bar{p}) \right] \right\}^k \mathbf{\Gamma}^k \quad (8)$$

Each element of the sum consists of two components. First,  $\mathbf{\Gamma}^k$  reflects the influence weights among the group after  $k$  iterations, that is, element  $i, j$  of  $\mathbf{\Gamma}$  to the  $k$ -th power gives the relative influence that the period  $t$  identity of group member  $j$  has on the period  $t + k$  identity of group member  $i$ . For  $k > 1$  this matrix also captures indirect influences and allows group members to influence each other that are not directly connected in the influence network. For example, with  $k = 2$ , element  $i, j$  of  $\mathbf{\Gamma}^k$  would be positive if  $j$  is at most 2 steps away from  $i$  in the influence network. The second component  $(\mathbf{I} - \mathbf{\Lambda}) \left[ \mathbf{I} - \hat{\mathbf{G}}(\bar{p}) \right]$  is a diagonal matrix with elements on the diagonal  $(1 - \lambda_i)\hat{g}(\bar{p}_i; \alpha_i) < 1$  that act as individual specific discount factors: the higher  $k$ , the smaller the factor that is multiplied with the term  $\mathbf{\Gamma}^k$ .

For each group member the expression above thus adds up the direct and indirect influences from across the group member network and weighs them in a way such that shorter network connections receive higher weight than longer ones. The infinite sum presents the relative influence weights in the limit when all possible influence paths through the group

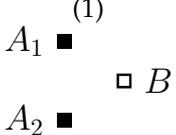
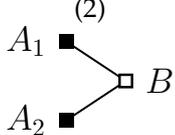
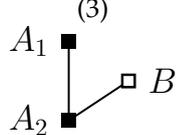
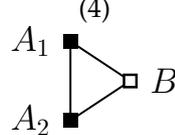
---

<sup>28</sup>Bonacich centrality is a recursive measure of node prominence in a network which assigns higher centrality to nodes that are connected to nodes that themselves have a high centrality. Formally, it is defined by  $C(\mathbf{\Gamma}, a, b) = \{\mathbf{I} - b\mathbf{\Gamma}\}^{-1} a\mathbf{\Gamma}\mathbf{1}$ , where  $a\mathbf{\Gamma}$  represents the so-called *base value* and  $b$  is a scalar that describes the *decay factor*, that is the rate at which the contribution of other nodes to node  $i$ 's centrality decays with their distance from  $i$ . For further details, see e.g. pp.40-43 in Jackson (2008). The weighted Bonacich centrality allows for heterogeneity in  $a$  across nodes (Ballester et al., 2006, p. 1409).

member network are accounted for. Overall, equation (8) can thus be seen as a matrix of influence weights that captures the long-run pattern of influence, once all direct and indirect influences amongst the group members have played out.

This matrix of long-run weights is then applied to the vector of idiosyncratic identity terms  $[\mathbf{I} - \hat{\mathbf{G}}(\bar{p})]\mathbf{A}$  from above, that captures the identity level of each group member without network influence. The result is a vector of steady state identities that each present a weighted average of the idiosyncratic identities, with the weights given by the infinite sum expression in equation (8).

To illustrate the relationship expressed in Proposition 4 between characteristics of the group network and the resulting long run identities consider the following simple example involving three group members. Two group members,  $A_1$  and  $A_2$ , have a high productivity  $\alpha_A$ , the third group member  $B$  a lower productivity  $\alpha_B$ , with  $\alpha_A > \alpha_B$ . We consider different network structures among the three. In each case, we focus on symmetric distributions of influence weights such that if one group member is connected to exactly  $n$  other group members, each neighbor and the group member himself receive weight  $1/(n + 1)$  in the updating process. Table 1 shows the results.

	(1)	(2)	(3)	(4)
				
Degree ( $A_1, A_2, B$ )	0, 0, 0	1, 1, 2	1, 2, 1	2, 2, 2
Clustering Coefficient	0	0	0	1
Density	0	2/3	2/3	1
Long-Run Identities ( $A_1, A_2, B$ )	 (0.20, 0.20, 0.92)	 (0.28, 0.28, 0.66)	 (0.21, 0.25, 0.70)	 (0.25, 0.25, 0.65)

**Note:** The example uses  $\hat{g}(p; \alpha) = \alpha e^{-0.9p}$ . With  $\alpha \leq 1$  this function ensures  $\hat{g}(p; \alpha) \in (0, 1)$  and in addition satisfies Assumption 1. Other parameters are as follows:  $p_L = 1$ , initial  $p_i^0 = 1 \forall i$ ,  $\alpha_A = 0.8$  and  $\alpha_B = 0.1$ .

Table 1: Long-Run Identities as a Function of Network Structure

(1) *Isolated Group Members* In the isolated case, long-run identities for the productive group members  $A_1$  and  $A_2$  converge to a common level that is significantly lower than that of group member  $B$ .

(2) / (3) *Open Triangle / Line* In the open triangle, we distinguish two cases. In the first case in

column (2),  $B$  is placed at the center, connecting to both  $A_1$  and  $A_2$ , while in the second case in column (3), one of the high productivity group members is located at the center. We see that in both cases, the existence of connections between the group members attenuates the differences between the high and low productivity group members relative to the isolated structure. In the case with  $B$  at the center,  $A_1$  and  $A_2$  show the same long run identity and  $B$ 's identity is lower than in the second case with  $B$  at an end node. In that second case,  $A_1$  and  $A_2$  will have different steady state identities, with the group member that is directly connected to  $B$  converging to an identity above that of the other group member.

(4) *Closed Triangle / Circle* In the case of the closed triangle,  $A_1$  and  $A_2$  once more display the same long run identity reflecting their symmetric positions in the group. Note that the extra connections attenuate the differences between high and low  $\alpha$  group members. These are smaller than in both the isolated and the open triangle setting.

The examples illustrate a number of features of the long run identities as characterized in Proposition 4. First, individual attributes such as the ability parameter  $\alpha_i$  and the individual specific leader influence  $\lambda_i$  are important drivers of differences in identity between group members. Second, connections in the influence network tend to move the long-run identities of the connected group members closer together.

In our example, the high productivity types  $A$  have a low identity when isolated but their identities become more pronounced when we add connections to the low productivity type, and vice versa for low productivity type  $B$ . In general, additional connections help attenuate the difference between the most extreme identities. A similar pattern emerges for clustering in groups. Comparing columns (2) and (4) in Table 1, we find that the difference between types  $A$  and  $B$  is smaller in the highly clustered closed triangle.<sup>29</sup>

In light of these results, we can also discuss the impact that removing individual group members has on the remainder of the group. As before, the impact of this intervention will depend on both idiosyncratic factors (such as ability and susceptibility to leader influence) and the extent of connectivity. First, a group member whose identity is more different from that of the rest of the group will have a larger effect than one who is more similar to others, holding all else equal. Second, if two group members have similar characteristics but vary in connectivity, then similar to the “key player” analysis of Ballester et al. (2006) the influence matrix also matters and the removal of a more connected player

---

<sup>29</sup>Note however that from this simple example we cannot disentangle the effect of clustering from the effect of the additional connection between  $A_1$  and  $A_2$ .

will have a greater impact than that of a less connected one.

Generally, the long run outcome of an individual group member depends on the interplay of both idiosyncratic factors and network structure. Holding social influence constant as in our example, higher ability types will integrate more and adopt a lower long-run identity than lower ability types. However, if the influence networks vary significantly between types, then it is possible for a higher ability group member to integrate less and have a higher identity in the long run than a lower ability group member.

**Comparative Statics of Ability  $\alpha$  and Group Division  $\gamma$**  If the steady state is unique, we can study the effect of  $\alpha_i$  as well as the effect of  $\gamma_{ij}$  on steady state identities as foundation for our analysis of policy instruments. To focus on the parameters of interest we assume  $\lambda_i = \lambda \forall i$ . If all group members have the same  $\alpha_i$ , that is the same productivity, then all group members will respond symmetrically and thus have the same  $\hat{g}(\bar{p}_i; \alpha_i)$  function. As a consequence the long run identities of the group members are the same, independently of the weights on other group members  $\gamma_{ij}$ .

**Remark 5** ( Ability  $\alpha_i = \alpha \forall i$ ).

*If all group members have the same level of ability and leader influence, the steady state identity vector is invariant to the structure of the group network.*

In contrast, if there is heterogeneity in ability  $\alpha_i$ , then the long run identities of the group members differ. To see this more clearly, suppose again that there are two types of group members,  $A$  and  $B$ . These two types differ in their ability, such that  $\alpha_A > \alpha_B$ . We furthermore assume that the influence structure is fully symmetric across members and types and captured by a single parameter  $\gamma$ , the group division specifying the total weight put on all group members of the same type.<sup>30</sup> Furthermore, both types share the same leader influence parameter  $\lambda_i$ . Steady state identity levels for  $A$  and  $B$  in this setting can then be summarized by:

$$\bar{p}_A = [1 - \hat{g}(\bar{p}_A; \alpha_A)] \{ \lambda p_L + (1 - \lambda) [\gamma \bar{p}_A + (1 - \gamma) \bar{p}_B] \} \quad (9)$$

$$\bar{p}_B = [1 - \hat{g}(\bar{p}_B; \alpha_B)] \{ \lambda p_L + (1 - \lambda) [\gamma \bar{p}_B + (1 - \gamma) \bar{p}_A] \} \quad (10)$$

Note that for a given identity level group member  $A$  invests more in skills than  $B$  resulting

---

<sup>30</sup>Note that as group members of the same type are homogeneous, the within-type influence weights do not affect the dynamics and final outcome.

in a higher weight on the host society for  $A$ . Based on equations (9) and (10) we can then analyze steady state behavior of this setting in Proposition 6.

**Proposition 6** (Identity Spread for Different Types). *Suppose there are two types of group members,  $A$  and  $B$ , with  $\alpha_A > \alpha_B$ . Then, in the steady state:*

1. *Type  $A$  agents assimilate more than those of type  $B$  that is  $\bar{p}_A < \bar{p}_B$ .*
2. *A higher group division, that is, a higher  $\gamma$  leads to more dispersed identities.*

The difference between steady state identities of  $A$  and  $B$  is decreasing in the weight they assign to each other, that is  $\frac{\partial \bar{p}_A}{\partial \gamma} < 0$  and  $\frac{\partial \bar{p}_B}{\partial \gamma} > 0$ . This implies that if group division is high, that is, group members do not influence each other, then the steady state identities differ greatly across the two types. On the other hand, if the community is more connected, then skill investment and identities are less dispersed and there is greater equality in outcomes. This also carries over to skill acquisition and earnings. If identities are more different, then so are earnings.

## 4 STRATEGIC LEADER IDENTITY CHOICE

Up to this point, the norms and the values of the leader have been taken as given. In this section we consider the strategic choice a leader faces and the conditions under which leaders implement different steady-state levels of assimilation of their community. In particular, we show when the leader implements an intermediate level of assimilation that falls between the full integration and the maximum distinction outcomes. We then proceed to an analysis of comparative statics and policy implications in Section 5.

In what follows we assume that group members are homogeneous with  $\alpha_i = \alpha$  and  $\lambda_i = \lambda$  for all  $i$ . In addition we simplify the marginal cost function of investment in skills such that  $c(p^t) = c_0 + c_1 p^t$  with  $c_0 > 0$  and  $c_1 > 0$ .

The leader's payoff is a function of both identity  $\bar{p}$  and group members' earnings  $\alpha f(\bar{H})$ . The steady state skill investment is again determined by the steady state identity and so we rewrite income as  $\alpha \hat{f}(\bar{p}; \alpha) \equiv \alpha f(H^*(\bar{p}; \alpha))$ . Note that  $\hat{f}(\bar{p}; \alpha)$  is *decreasing* in  $\bar{p}$ . This derives from the fact that  $f(\bar{H})$  is increasing in  $\bar{H}$  and  $H^*(\bar{p}; \alpha)$  is decreasing in  $\bar{p}$ .

The leader sets  $p_L$ , which influences  $\bar{p}$  through the steady state condition:

$$\bar{p} = [1 - \hat{g}(\bar{p}; \alpha)] \{ \lambda p_L + (1 - \lambda) \bar{p} \}, \quad (11)$$

Given Assumption 1, the impact of  $p_L$  on  $\bar{p}$  is strictly positive.<sup>31</sup> This implies that a higher identity of the leader is always associated with lower assimilation of group members.

We focus in what follows on the case where the leader chooses intermediate assimilation ( $0 < p_L < p^{\max}$ ) for the general payoff function  $\pi(\bar{p}, \alpha \hat{f}(\bar{p}, \alpha))$  (instead of extremism ( $p_L = p^{\max}$ ) and full assimilation ( $p_L = 0$ )).

**Proposition 7** (Intermediate Assimilation). *If there exists an interior solution and  $\hat{f}(\bar{p}; \alpha)$  is a concave function, then the leader implements an intermediate level of assimilation with  $p_L \in (0, p^{\max})$ .*

If an interior solution exists and the group member's payoff function is a concave function in  $\bar{p}$ , then intermediate cultural distinction emerges.<sup>32</sup> A concave  $\hat{f}(p; \alpha)$  implies that a small decrease in identity at  $p^{\max}$  leads to a large increase in investment in skills and thus to a large increase in earnings. On the other hand, an increase in identity at  $p_L = 0$  leads to a negligible loss in earnings. This implies that the leader can increase the community's income drastically by reducing cultural distinction if identity is high, but there is no great loss in earnings if he increases cultural distinction if it is low. Thus, with a concave earnings function the leader will prefer the community to assimilate to some extent. Note that the concavity of  $\hat{f}(\bar{p}; \alpha)$  is a sufficient condition. Thus, if the earnings function  $\hat{f}(p; \alpha)$  is convex and an interior solution exists, then it might still be optimal for the leader to choose intermediate assimilation.

Figure 2 illustrates the leader's problem and conveys some of the intuition behind Proposition 7. The thicker line shows the feasible combinations of group member earnings and group member identity that can be implemented with a given level of leader identity  $p_L$ . The feasible group member identities lie between 0 and  $\bar{p}(p^{\max})$ , that is, the group member identity that results from the leader selecting the largest possible degree of distinction. Note that even at that level, there is some skills investment by group members and thus positive earnings. The thinner curves are indifference curves of the leader payoff function. The optimal leader identity selects the point on the thicker line that lies on the highest possible indifference curve.

In order to provide some more intuition on possible assimilation outcomes, we turn to the special cases of the religious and business leader.

<sup>31</sup>See the proof to Proposition 7 in the Appendix for a formal proof.

<sup>32</sup>The concavity of the earnings function is driven by the sign of  $f'''(H)$ , with  $f'''(H) < 0$  guaranteeing a concave earnings function. We provide further details in the Appendix.

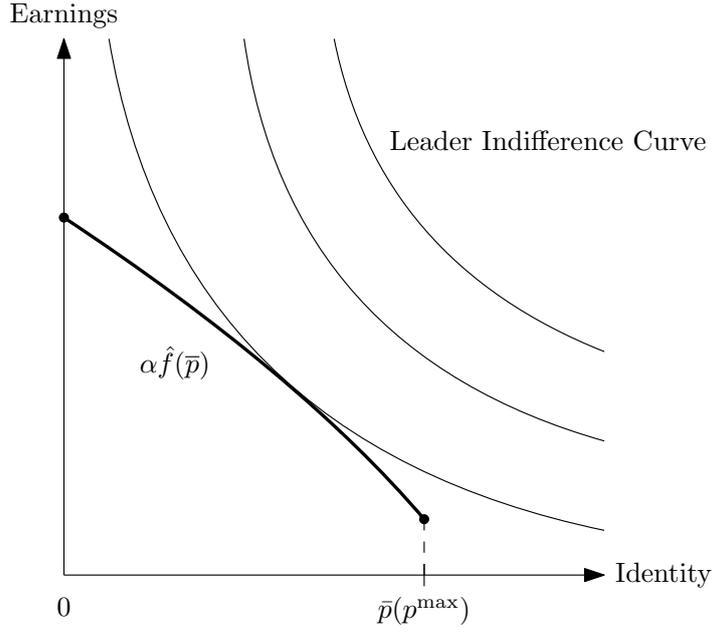


Figure 2: Illustration of Strategic Leader Identity Choice

**Proposition 8** (Religious & Business Leader). .

1. Unless there exists an interior solution and  $\hat{f}(\bar{p}; \alpha)$  is a concave function, the religious leader selects either the most extreme norms and values,  $p^{\max}$ , or full assimilation.
2. The business leader never chooses full assimilation.

In the case of the religious leader, concavity of  $\hat{f}(\bar{p}; \alpha)$  ensures intermediate assimilation, which follows immediately from Proposition 7. However, it can also be the case that the religious leader chooses extremism or full assimilation. The latter occurs if the earnings function is convex as this implies that a small decrease in  $p$  away from  $p^{\max}$  has only a small impact. In order to realize high earnings with a convex  $\hat{f}(p; \alpha)$ , the identity has to decrease significantly. Therefore, if the earnings function is convex, the religious leader chooses either full assimilation or no assimilation at all, depending on which extreme identity yields the highest payoff.<sup>33</sup>

Similar to the religious leader, the business leader might also choose the most extreme level of norms, if he does not opt for intermediate assimilation. However, unlike the religious leader, it is never optimal for the business leader to set  $p_L = 0$ . If he were to do so, he would earn zero profits. A small increase in identity increases earnings and leads to strictly positive profits.

<sup>33</sup> This depends on  $p^{\max}$  as well as on  $\alpha$  and  $\beta$ . If  $p^{\max}$  is sufficiently large, then the religious leader always prefers the highest possible cultural distinction. For smaller  $p^{\max}$ , a higher productivity  $\alpha$  and lower  $\beta$ , where  $\beta$  is the weight the leader assigns to identity, make full assimilation more likely.

If group members are only willing to invest in skills if the payoff from this is sufficiently high, see again equation (5), then the leader's maximization problem is unchanged. If the optimal outcome is extremism or intermediate assimilation with group members investing in skills, then the result remains unchanged. Complete assimilation can no longer be an equilibrium, which follows trivially from the fact that lowering identity below a certain threshold will not result in an increase in skills and therefore earnings. The logic is similar for intermediate assimilation with group members not investing in skills. This induces the leader to set a higher identity, which results in less assimilation.

## 5 Comparative Statics and Policy Implications

We now turn to the comparative statics and connect our findings to policy implications. We consider in turn the effect of a change in productivity  $\alpha$ , the effect of a lump sum transfer,  $z$ , and last the effect of a change in group ties  $\gamma$ .

**Effect of Change in Productivity** We are interested in the effect of an increase in  $\alpha$ , that is an increase in productivity. Such an increase in productivity can emerge through lower discrimination, as with lower discrimination investment in skills leads to a higher payoff. In this sense considering the effect of  $\alpha$  sheds light on the impact of a policy aimed at decreasing discrimination.<sup>34</sup> Our finding is summarized in Proposition 9.

**Proposition 9** (Comparative Statics  $\alpha$ ). .

*If the leader chooses intermediate assimilation, then as  $\alpha$  increases:*

1. *The leader selects a higher  $p_L$  resulting in higher group member identity  $\bar{p}$ .*
2. *Group member income increases.*

In case of intermediate assimilation, a higher  $\alpha$  results in an increase in the identity the leader sets. An increase in  $\alpha$  leads to an increase in the marginal payoff from investing in skills, resulting in an increase in skill investment for a given leader identity. In response, the leader adjusts his identity and in fact he will increase it, resulting in both wealthier and less assimilated group members. Note that identity and skill investment are not only affected

---

<sup>34</sup>Recall that skill investment is only beneficial for individuals when they enter the native labor market. Skills should not be thought of as, for example, education but rather language skills and understanding of norms and values in the host society. Therefore, if the returns to these skills are low, then it can be interpreted as facing a host country with high discrimination.

by marginal changes, but also by level effects, which will become clear when we consider the effect of a lump sum transfer.

To shed some further light on when skills increase, remain constant or decrease we consider the steady state payoffs of the religious and business leaders.

**Corollary 9.1** (Religious vs Business Leader: Skill Investment). .

*With a religious leader skill investment is unaffected by a change in  $\alpha$ , whereas with a business leader skill investment increases.*

To see that skill investment remains constant in case of a religious leader, consider the first order condition given by

$$\beta + (1 - \beta)\alpha f'(\bar{H}) \frac{\partial \bar{H}}{\partial \bar{p}} = 0 \quad \Leftrightarrow \quad \beta + (1 - \beta) \frac{f'(\bar{H})}{f''(\bar{H})} c_1 = 0 \quad (12)$$

Equation (12) is independent of  $\alpha$  and so the optimal  $H^*$  is unaffected by a change in  $\alpha$ . In contrast, for the business leader a change in  $\alpha$  increases skill investment. The identity is not increased by so much as to crowd out the increase in skill investment completely.

To summarize, an increase in productivity  $\alpha$  without the presence of a leader induces immigrants to invest more in skills, which results in higher wages and greater assimilation and, thus, higher cultural distinction is related to lower earnings. In the presence of a leader this result does no longer hold as the leader sets a higher level of cultural distinction which leads the community to be less integrated. The increase in  $\alpha$  thus leads to both higher wealth and higher identity.

**Effect of a Lump Sum Transfer** We are interested in the effect of a lump sum transfer  $z$  to the immigrant group member from the host country on identity, earnings and skill investment. Many European countries have social security systems that are also accessible for immigrants. We are interested in the effect of these security systems on cultural integration. The income of a group member who obtains these transfers is then given by  $\alpha \hat{f}(\bar{p}; \alpha) + z$ .

**Proposition 10** (Effect of a Lump Sum Transfer). *Suppose the leader chooses intermediate assimilation. Then, a lump sum transfer  $z$  leads to weakly higher cultural distinction, weakly lower skill investment and higher earnings.*

An exogenous increase in earnings that does not hinge on skill investment can change the decision problem of the leader. In particular, it might decrease skill investment as the

leader is not dependent on this investment to increase the community's earnings. This implies that a group member's identity might increase. Last, the effect on overall earnings is positive. As we have seen previously, a religious leader chooses always the same amount of skill investment. Therefore, in this case, a transfer has no effect. This is not true for the business leader, where the lump sum transfer will increase identity and decrease skill investment.

**Remark 11** (Effect of a Lump Sum Transfer: Business Leader vs Religious Leader). *A business leader choosing intermediate assimilation will increase the identity selected in response to an increase in  $z$ . This in turn leads to lower skill investment of the group members. In case of a religious leader, a lump sum transfer has no impact on the identity selected. This holds if the religious leader chooses intermediate assimilation as well as if he decides between full assimilation and extremism.*

This shows that a religious leader's position is never affected by a transfer. In particular, a religious leader who chooses extremism is not more inclined to switch to the full assimilation equilibrium. Denote by  $\bar{p}(p^{\max})$  the identity reached when the leader sets  $p_L = p^{\max}$ . Suppose the leader sets the maximum identity and prefers this to full assimilation. Then, he will also be an extremist in case of a transfer as

$$\begin{aligned} & \beta \bar{p}(p^{\max}) + (1 - \beta)(\alpha \hat{f}(\bar{p}(p^{\max})) + z) > (1 - \beta)(\alpha \hat{f}(0) + z) \\ \Leftrightarrow & \quad \beta \bar{p}(p^{\max}) + (1 - \beta)(\alpha \hat{f}(\bar{p}(p^{\max}))) > (1 - \beta)\alpha \hat{f}(0). \end{aligned}$$

Our finding sheds light on the fact that in our setting it is not just a change in the marginal productivity that has an effect on skill investment, identity and earnings, but also a level change.

Similar to the effects of an increase in  $\alpha$ , an increase in the lump sum transfer  $z$  leads to a weakly higher identity and strictly higher earnings.

We abstract here from an analysis of the impact of an increase in productivity  $\alpha$  or a lump sum transfer  $z$  on welfare. It is natural to assume that welfare depends both on earnings and identity.<sup>35</sup> Then an increase in  $\alpha$  makes individuals better off (assuming costs are not too high), as *both* earnings and identity increase following an increase in either  $\alpha$ . Similarly, an increase in  $z$  makes individuals better off. Thus, these type of policies increase the

---

<sup>35</sup>Alternatively, welfare could only depend on earnings and then the effect of a parameter change on welfare is the same as that on earnings.

welfare of immigrants, but they do not foster cultural integration.<sup>36</sup> This is summarized in the following remark.

**Remark 12.** *Higher earnings do not affect the degree of integration.*

We can have immigrant groups that have a high cultural distinction from the host society, but as they are very productive, they still have high earnings. In this case, earnings and identity are positively correlated. However, this comparison only holds across groups. Within groups agents with higher productivity have higher earnings than those with lower productivity and so earnings and identity are negatively related.

**Heterogeneous Group Members** Next, we allow again for heterogeneity in ability. We restrict attention to two types,  $A$  and  $B$  with  $\alpha_A > \alpha_B$ , and the symmetric influence structure summarized by equations (9) and (10). We ask how the optimal identity set by the leader responds to group members being more or less connected, that is how changes in the measure of group division  $\gamma$  affects  $p_L$ . Recall that a lower  $\gamma$  implies that an individual is more strongly influenced by his peers of another type. Our analysis sheds some light on the fact that groups in which ties are more important exhibit a very different assimilation pattern, not only because of the ties per se, but also because of the effect on the leader. This can have effects on how assimilationist policies might be structured for different types of groups.

**Proposition 13** (Comparative Statics  $\gamma$ ). *Consider the two-type setting with  $\alpha_A > \alpha_B$  characterized in equations (9) and (10). Let  $\hat{g}(\bar{p}_i; \alpha_i)$  be a concave function and  $\gamma \geq \frac{1}{2}$ . Additionally, assume that  $\frac{\partial \pi(\bar{p}, \alpha \hat{f}(\bar{p}; \alpha))}{\partial \bar{p}}$  is concave. Suppose the leader chooses intermediate assimilation. Then, a decrease in group division  $\gamma$  leads to a greater level of cultural distinction, to a higher  $p_L$ . The identity of  $A$  increases and so his investment and earnings decrease. The effect on identity and thus investment and earnings of  $B$  is ambiguous.*

One implication of  $\hat{g}(\bar{p}_i; \alpha_i)$  being concave is that an increase in the weight that group members assign to each other will lead to an increase in  $\bar{p}_A$  that is smaller than the decrease  $\bar{p}_B$ , keeping  $p_L$  fixed. However, the leader has an incentive to adjust his identity and in particular, if  $\gamma$  decreases then the leader has an incentive to increase  $p_L$ . To see this more clearly, consider the first order condition of the leader when faced with a group that consists

---

<sup>36</sup>It is straightforward to incorporate an explicit welfare analysis as the impact of the  $\alpha$  and  $z$  on earnings and identity go in the same direction.

of two types:

$$\left( \frac{\partial \pi(\bar{p}_A, \bar{w}_A)}{\partial \bar{p}_A} + \frac{\partial \pi(\bar{p}_A, \bar{w}_A)}{\partial \bar{w}_A} \frac{\partial \bar{w}_A}{\partial \bar{p}_A} \right) \frac{\partial \bar{p}_A}{\partial p_L} + \left( \frac{\partial \pi(\bar{p}_B, \bar{w}_B)}{\partial \bar{p}_B} + \frac{\partial \pi(\bar{p}_B, \bar{w}_B)}{\partial \bar{w}_B} \frac{\partial \bar{w}_B}{\partial \bar{p}_B} \right) \frac{\partial \bar{p}_B}{\partial p_L} = 0 \quad (13)$$

Given our assumption that induces an interior solution, we know that

$$\left( \frac{\partial \pi(\bar{p}_A, \bar{w}_A)}{\partial \bar{p}_A} + \frac{\partial \pi(\bar{p}_A, \bar{w}_A)}{\partial \bar{w}_A} \frac{\partial \bar{w}_A}{\partial \bar{p}_A} \right) > 0 > \left( \frac{\partial \pi(\bar{p}_B, \bar{w}_B)}{\partial \bar{p}_B} + \frac{\partial \pi(\bar{p}_B, \bar{w}_B)}{\partial \bar{w}_B} \frac{\partial \bar{w}_B}{\partial \bar{p}_B} \right) \quad (14)$$

This implies that when  $\gamma$  decreases, for fixed  $p_L$ , the value of both the LHS and RHS of equation (14) move closer to zero. Additionally, as the leader has a greater influence on type  $B$ , it has to hold that  $\left( \frac{\partial \pi(\bar{p}_A, \bar{w}_A)}{\partial \bar{p}_A} + \frac{\partial \pi(\bar{p}_A, \bar{w}_A)}{\partial \bar{w}_A} \frac{\partial \bar{w}_A}{\partial \bar{p}_A} \right) > \left| \left( \frac{\partial \pi(\bar{p}_B, \bar{w}_B)}{\partial \bar{p}_B} + \frac{\partial \pi(\bar{p}_B, \bar{w}_B)}{\partial \bar{w}_B} \frac{\partial \bar{w}_B}{\partial \bar{p}_B} \right) \right|$  for equation (13) to hold. This implies that the decrease in  $\gamma$  leads equation (13) to be positive and it is optimal for the leader to increase  $p_L$ . Therefore, societies with a lower group division, that is, higher group cohesion, will induce the leader to choose a higher level of cultural distinction.

Note that the assumption that  $\frac{\partial \pi(\bar{p}, \alpha \hat{f}(\bar{p}; \alpha))}{\partial \bar{p}}$  is a concave function is a sufficient condition and our results also hold as long as the difference between the second order conditions,

$$\frac{\partial^2 \pi(\bar{p}_A, \alpha_A \hat{f}(\bar{p}_A; \alpha_A))}{\partial \bar{p}_A^2} \quad \text{and} \quad \frac{\partial^2 \pi(\bar{p}_B, \alpha_B \hat{f}(\bar{p}_B; \alpha_B))}{\partial \bar{p}_B^2}$$

of group members  $A$  and  $B$  is not too large.<sup>37</sup>

The increase in  $p_L$  following a decrease in  $\gamma$  leads for group member  $A$  to an increase in identity. His identity increases when  $\gamma$  decreases and  $p_L$  remains fixed. Now,  $p_L$  increases as well leading to a further rise in his identity. From this it follows that the skill investment and earnings of the high productivity group member decrease. The effect on the identity of the low productivity group member is ambiguous. His identity would decrease if  $\gamma$  decreases and  $p_L$  were kept fixed. The decrease in  $p_L$  now goes against this effect and so the overall effect on identity is ambiguous. The same holds true for the effect on his skill investment and earnings.

<sup>37</sup> The condition is always fulfilled for the religious leader if  $f^{(4)}(H) \leq 0$ .

## 6 Cutting Ties

We have assumed throughout that group members' connections to their own community are weakened as they become more integrated in the host society. However, in our baseline setting group members never fully leave their group and it seems plausible that agents can also cut their connection to the community once their identity, that is their identification with their own group, is sufficiently low. We consider this scenario, in which agents sever ties to their community if they are sufficiently integrated, in what follows.

Note first that in the setting with two groups, there are no incentives for the group leader to unilaterally exclude one group from the community.

**Proposition 14.** *In the two-type setting with  $p^{\max}$  sufficiently high, the leader never excludes one group from the community.*

However, group members might still leave according to their own wishes. Suppose there exists a boundary identity,  $\underline{p}$ , such that all group members with an identity lower than  $\underline{p}$  sever ties to the group and only members with a higher identity remain in the community. How does the leader respond to the possibility of group members cutting ties?

The problem is trivial if there is no group member whose identity lies below the bound at the optimum: the leader is then effectively unconstrained and simply behaves as if the bound were not present. However, if a group member's identity at the unconstrained maximum lies below the bound, which induces this group member to leave the community, the leader has two possibilities to respond. First, he can accommodate the group members below the bound by increasing his own identity which in turn increases the identities of the group members until no group member has the incentive to leave the community. Alternatively, he allows for the group members to leave and focuses on the remaining group members, selecting a possibly different level of  $p_L$  that is optimal for the smaller group. In order to calculate which option is more beneficial, we need to compare the payoff from excluding a group to the payoff from adjusting the identity to retain the group.

To illustrate this comparison, we focus on the case of two types,  $A$  and  $B$ , with symmetric influence pattern where as before we assume that group members of type  $A$  have a higher ability than group members of type  $B$ , such that in steady state  $\bar{p}_A < \underline{p} < \bar{p}_B$ . Denote the share of group members of type  $A$  by  $\kappa$ . We index variables with a superscript  $B$  to indicate that only members of type  $B$  belong to the group, and by  $AB$  if both types remain

in the group. Then,  $\Pi^B$ , that is, the payoff if only type  $B$  remains in the group, is given by

$$\Pi^B = (1 - \kappa) \pi(\bar{p}_B^B, \alpha_B f(\bar{H}_B^B))$$

If both types remain in the group, then  $\bar{p}_A^{AB} = \underline{p}$ . We denote  $B$ 's identity in this case by  $\bar{p}_B^{AB}$ . Then, the payoff is given by

$$\Pi^{AB} = \kappa \pi(\underline{p}, \alpha_A f(\bar{H})) + (1 - \kappa) \pi(\bar{p}_B^{AB}, \alpha_B f(\bar{H}_B^{AB}))$$

It is straightforward to see that the smaller  $\kappa$ , the more likely it is that the leader does not increase his identity to retain members of type  $A$ . Depending on the exact specification it can either be optimal for the leader to retain  $A$  types or to let them go. We construct an example for each of these outcomes for a religious leader with the functional forms shown in Table 2.

$$\frac{f(H)}{kH - \frac{1}{6}H^3} \quad \frac{g(H)}{\frac{H}{1+H}} \quad \frac{c(p)}{\frac{1}{2}(1+p)}$$

Table 2: Functional Forms

Then, we can provide an example for which  $\Pi^B < \Pi^{AB}$  and one for which the reverse holds. It is better to keep both group members in the community,  $\Pi^B < \Pi^{AB}$  for the parameters specified in the first row of Table 3, whereas the second row shows the parameters under which it is optimal to omit type  $A$ . Based on these parameters we can show what the

	$\kappa$	$\beta$	$\lambda$	$\gamma$	$\alpha_A$	$\alpha_B$	$k$	$\underline{p}$
(1) $\Pi^B < \Pi^{AB}$	1/2	1/4	1/2	1/2	3/4	1/2	4	3/4
(2) $\Pi^B > \Pi^{AB}$	1/200	2/5	1/2	1/2	2	1/2	3	9/10

Table 3: Example

steady state identities, the identities of the leader and earnings of the group members are, respectively. This is summarized in Table 4, where again the first row gives the case when it is better to keep group member  $A$  within the community, whereas in the second row it is better if  $A$  has severed his ties.

Note that in the first case where it is optimal to retain group  $A$ , the leader would prefer to set a lower identity than 3.93 if he only faces  $B$  types. However, he cannot do so as then the group members of type  $B$  would leave the group. What can be noted here is

	$p_L^B$	$\bar{p}_B^B$	$\alpha_B f(\bar{H}_B^B)$	$\Pi^B$	$p_L^{AB}$	$\bar{p}_B^{AB}$	$\alpha_A f(\bar{H}_A^{AB})$	$\alpha_B f(\bar{H}_B^{AB})$	$\Pi^{AB}$
(1)	3.93	.75	3.45	1.38	4.29	.82	5.46	3.42	3.52
(2)	3.70	.95	1.92	1.52	4.70	1.39	9.70	1.55	1.50

Table 4: Identities and Earnings

that whenever only type  $B$ 's remain in the group, the leader sets a lower identity, which results in better integration and higher earnings of the group members. Therefore,  $A$  types leaving the group leads to better integration outcomes through two channels: the  $A$  types will integrate completely and additionally, the  $B$  types will integrate better.

It is natural to ask to what extent the effect of a policy measure in the case where ties vary only in their intensity carry over to the setting where ties can be cut. To see this more clearly, suppose  $B$ 's productivity  $\alpha_B$  decreases, from  $1/2$  to  $1/3$ . We focus here on example (1) as such a decrease in  $\alpha_B$  in example (2) induces the leader to choose maximal cultural distinction (independently of whether the boundaries are present or not). The identities and earnings if ties only change continuously are given in the first row of Table 5. The second row shows the corresponding value if  $\alpha_B$  decreases from  $1/2$  to  $1/3$ .

	$p_L$	$\bar{p}_A$	$\bar{p}_B$	$\alpha_A f(\bar{H}_A)$	$\alpha_B f(\bar{H}_B)$	$\Pi$
(1) with $\alpha_B = 1/2$	4	.70	.76	5.47	3.44	3.52
(1) with $\alpha_B = 1/3$	1.45	.24	.28	5.56	2.25	2.99

Table 5: Identities and Earnings without Cutting Ties

This shows that if  $\alpha_B$  decreases, the leader lowers his identity significantly. Both types of identities fall and while the earnings of type  $A$  increase, those of type  $B$  decrease.

If ties to the community are cut for sufficiently low identity, then the payoffs for  $\alpha_B = 1/3$  are as given in Table 6.

	$p_L^B$	$\bar{p}_B^B$	$\alpha_B f(\bar{H}_B^B)$	$\Pi^B$	$p_L^{AB}$	$\bar{p}_B^{AB}$	$\alpha_A f(\bar{H}_A^{AB})$	$\alpha_B f(\bar{H}_B^{AB})$	$\Pi^{AB}$
(1)	3.93	.75	1.96	.83	4.15	1.10	5.46	1.61	2.88

Table 6: Identities and Earnings:  $\alpha_B = 1/3$

If only type  $B$  remains in the group, the leader has to set the same identity as before the change in  $\alpha_B$  in order to satisfy the boundary condition. If both group members remain in the group, then the leader sets a lower identity similar to the case where ties cannot be cut. Given that group members  $B$  now have a higher identity for a given leader choice  $p_L$ , the

leader can reduce the  $p_L$  he chooses in order to keep  $A$  within the group. Therefore, the policy implications in this case are exactly the same as if ties cannot be cut.

Generally, we consider the case in which all group members remain in the group to be the most interesting and realistic one. As can be seen from the example presented, the share of productive types needs to be very small in order for the productive group members to be omitted from the group.<sup>38</sup> If it is only beneficial to let group members leave, if they constitute a sufficiently small part, then policy makers are still faced with large groups of immigrants who will not integrate fully and essentially the problem remains the same as in the case where ties are not cut. In fact the problem is exacerbated, as group leaders will now choose a higher level of cultural distinction compared to the case where ties cannot be cut and the impact of the policies is more limited. Whereas the leader's identity fell dramatically when cutting links was not possible as a response to the decrease in  $\alpha_B$ , the reduction in the leader's identity if ties can be eliminated is much more moderate.

## 7 DISCUSSION

The predictions we derived can be connected to several stylized facts and can help understand different patterns in the assimilation of immigrants or minorities as well as possible policy implications.

**Jewish Assimilation after Emancipation** The case of the Jewish Assimilation after Emancipation in Germany and Hungary is documented in [Carvalho and Koyama \(2011\)](#). They show that Jewish communities in Germany assimilated to German customs under the influence of religious leaders. Rabbis changed the religious procedures such that less time was required to be spent at the synagogue. Additionally, organ music, traditionally a Christian custom, was introduced and sermons started to be held in German instead of Hebrew. All these changes can be seen as increased assimilation to German customs and traditions. However, the assimilation took place at the same time as the building of many new synagogues. During that period, the German communities became very wealthy and invested in the construction of many new synagogues.<sup>39</sup> In contrast, the community leaders in Hungary started creating norms and customs that required their followers to spend more time at the synagogue. They also imposed dress codes that emphasized the difference between the

---

<sup>38</sup> This seems to be a general feature of the examples we considered.

<sup>39</sup>Note that we are only interested in assimilation as far as it concerns economic outcomes. So, even though the Jewish community persisted it was on many important issues not that different from Christian communities.

Jewish community and the wider Hungarian society. [Carvalho and Koyama \(2011\)](#) argue that these difference in development stem from diverse economic environments: Germany was already industrialized, while the Hungarian economy was still largely agrarian.

In terms of our model, we think of rabbis as religious leaders that end up choosing between full assimilation versus extremism. In our model this requires that their earnings function is convex in identity. We argue that at the time, such an earnings function seems indeed plausible. A convex earnings function implies that a small change in identity, away from the community's original identity has a relatively small impact on earnings. In the case of the Jewish communities, this means that as long as individuals were still showing that they were Jewish to a noticeable degree they were not rewarded in terms of higher economic earnings. Only complete assimilation resulted in full admission to the economy. [Katz \(1980\)](#) offers supporting evidence for this. First, the process of emancipation was a slow one. It started out in 1780, but did not last in some places. In fact, there are instances where citizen rights were withdrawn again. Only by 1848-1849 did Jews have full equality in all of Germany. This provides evidence that Jews were not welcomed to be part of German society as Jews. Additionally, this was the period where strong anti-Semitism emerged in Germany, see [Katz \(1980\)](#).<sup>40</sup> Making cultural traits unnoticeable might therefore have helped to overcome these prejudices. As already emphasized, one key difference between Hungary and Germany was the productivity in the environment. Assimilation in Hungary would not have reaped the same benefits as in Germany. In terms of our model, this would imply significant differences in  $\alpha$  that can result in markedly different responses by leaders with a convex earnings function: full assimilation in one case and maximum differentiation in the other.

Our model thus sheds light on why different communities in different countries choose different paths for assimilation and how this is driven by their religious leaders. The example of the Jewish communities is additionally of interest, as these communities had exactly one religious leader, without competition: communities often did not agree with the leader's position but still had no alternative but to follow it. Thus, this is a setting that is very much in line with our model.<sup>41</sup>

### **Muslim Integration is Different from that of other Immigrants to Europe** Researchers

---

<sup>40</sup>The Jewish communities in Germany adapted to German cultural traits, despite great animosity towards them, which likely resulted in strong discrimination, a pattern that is surprisingly in line with the predictions of our model.

<sup>41</sup>For evidence on this and a summary of the situation of the Jewish communities, see [Carvalho and Koyama \(2011\)](#).

frequently comment that Muslim integration in Europe seems to differ from that of other communities (Bisin et al. (2008) for the UK; Constant et al. (2006) and Haug (2008) for Germany). The predictions of our model are in line with several features of this integration process. One possible reason for this is that Imams in Muslim communities seem to be of greater importance than religious leaders in other communities (see Ceylan (2010)), but in addition they also seem to differ along various other dimensions.

First, a majority of the Imams in Germany are employed by the Turkish government and serve for only 4-5 years in Germany. Their knowledge of German is limited, with only 50% of the Imams taking German classes (Halm et al. (2012)). Sermons are traditionally held in Turkish. The situation is similar in the UK, as documented in Geaves (2008) where more than 90% of Imams come from abroad. In addition, similar to the situation in Germany, the majority of Imams only arrived in the UK less than five years ago and was educated and raised in an environment with very different norms and cultural values than those prevalent in the UK. There are in fact a few Imams that were raised and educated in Britain. Nonetheless, even those Imams give sermons in Urdu, the prevalent language of origin for Muslims in the UK.

Second, the income of German Imams, due to their special employment status, does not depend on how wealthy their community is. This is different from Imams in the UK. However, it seems plausible that Imams put less emphasis on high earnings, given that they tend to come from a poor background. This seems also to be reflected in the pattern of earnings of religious leaders with Imams earning only little. The average earnings of clergy in the UK lies at approximately £20,000 (see the 2013 Annual Survey of Hours and Earnings), whereas it is reported that Imams make between £10,000 and £12,000 a year.<sup>42</sup> In terms of our model this translates into a lower  $\beta$ , which would lead to greater emphasis on cultural values compared to earnings and thus to a culturally more distinct community.

Additionally, many Muslim communities have strong family values. According to a standard classification based on the work of Hofstede (1984), Turkey and Pakistan, the two most prominent source countries for Muslim immigration to Germany and the UK, are defined to be collectivist, where an individual is less important compared to the community. In such an environment, we would argue that community members put a higher weight on the influence of other members. In terms of our model, this implies a higher level of cultural

---

<sup>42</sup>See an article in The Guardian on [www.theguardian.com/commentisfree/2007/jun/05/vacancyforanimam](http://www.theguardian.com/commentisfree/2007/jun/05/vacancyforanimam) (accessed 27.10.2016).

distinction, not only because the group is cohesive per se, but also because it induces the leader to impose values and norms of higher cultural distinction.

Another feature of the assimilation process is that Muslim immigrants tend to hang on to their cultural traits and religion more so than other immigrants. In particular, the assimilation of second generation immigrants is lower than that of other religious and ethnic groups. Arguably, Muslim immigrants come from a culturally more different background than other immigrant groups, slowing their assimilation process. This is in line with our prediction that assimilation occurs more quickly for those from an already similar background.

Further, unlike other groups, Muslims with higher wages tend to be more religious and more identified with their religion, compared to poorer Muslims. This is in line with our prediction that if leaders can set norms and values, then when comparing two distinct groups, as suggested in [Bisin et al. \(2008\)](#), the members of the group with higher ability end up with higher identity and higher earnings.

Finally, our model draws attention to the finding that Imams educated in the UK do not choose different norms and values compared to Imams coming from abroad ([Geaves \(2008\)](#)). We predict that leaders that are strategic will respond to the incentives that are in place in the community and thus different leaders may well end up with the same level of cultural distinction, independently of their individual background.

**Extremism and Wealth** Our model also sheds some light on the fact that higher cultural distinction is not correlated with wealth or earnings. Our findings are in line with [Krueger \(2008\)](#), who argues that poverty will not lead to the emergence of extremism and finds little correlation between extremism, terrorism and poor economic conditions.<sup>43</sup>

We consider two different channels of why wealth is not a predictor of cultural distinction. In particular, we distinguish between an increase in productivity and the effect of a lump-sum transfer. In case of a religious leader, who decides between full assimilation and extremism, we find that an increase in productivity might induce full assimilation. But this

---

<sup>43</sup>[Krueger \(2008\)](#) also shows that lack of education is uncorrelated with extremism. In particular, he emphasizes that increasing educational spending might even be counterproductive. This can be related to reducing the cost of skill acquisition, that is reducing  $c_0$  or  $c_1$  and can be interpreted as support for language courses. We have not spelled out the effect of such a policy in detail, mainly because the implications are very similar to that of a change in productivity and the introduction of a lump sum transfer. In particular, a decrease in  $c_1$  will lead to a higher cultural distinction if the leader chooses intermediate assimilation and is thus symmetric to an increase in productivity  $\alpha$ . A decrease in  $c_0$  is akin to a lump-sum transfer and will have a weakly positive effect on cultural distinction. With a religious leader it has again no effect and with a business leader it leads to higher cultural distinction. These results and proofs follow along the same lines as the proof of an increase in productivity  $\alpha$  and are omitted here.

is only the case if the jump in productivity is sufficiently high. It might also be that an increase in productivity has no effect whatsoever. In contrast, a transfer will never induce the religious leader to change from extremism to full assimilation. Generally, a transfer only affects earnings, and has no impact on identity or skills and thereby earnings. An increase in productivity, however, induces higher cultural distinction if the religious leader chooses intermediate assimilation. Higher productivity will lead to higher earnings without affecting skill investment.

Based on this we predict that when governments support immigrants economically and accommodate their specific cultural traits, then this will not foster cultural integration. This pattern has, in fact, been empirically documented by [Koopmans \(2010\)](#). He compares socio-economic outcomes of immigrants across different European countries and finds that countries that do not provide strong incentives for host-country language acquisition, but have a strong welfare state have produced low levels of labor market participation of immigrants. On the other hand, the United Kingdom, despite having multicultural policies in place, has a higher labor market participation. [Koopmans \(2010\)](#) argues that this is due to the relatively lean welfare state and suggests that governments might want to restrict access to welfare and social security in order to encourage immigrant labor force participation. Our analysis highlights that the instruments of a welfare state, even though they improve the welfare of immigrants, reduce the incentives to cultural integration.<sup>44</sup>

## 8 Conclusion

We develop a model of the assimilation process of an immigrant community with a leader, describing the evolution of the skills and identities of community members. In the absence of a leader, our setting predicts complete assimilation. The presence of a leader can prevent this and lead to persistent identity differences between immigrant community and host society. We characterize the circumstances in which leaders choose full assimilation, intermediate assimilation or extremism with maximum differentiation.

We proceed to show that higher productivity of the environment can lead to a community that is both wealthier and more identified with the immigrant community as the leader might choose higher cultural distinction. Further, we consider the effect of group members

---

<sup>44</sup>Targeting the income of the leader, for example by direct funding, likewise does not have the desired effect: by increasing the financial income of leaders the marginal return for wealth relative to identity decreases, once again providing an incentive for leaders to implement a lower degree of group assimilation.

being more influenced by their peers. If the community is more cohesive, then the leader can have incentives to choose more different norms and values.

Our model has been silent on the emergence of leaders. Many immigrants are second generation immigrants, who were born into a community in which they were exposed to a leader, rendering plausible the assumption of immigrants not selecting an their leader. Nonetheless, we believe that modeling the emergence of these leaders is a fruitful area for future research. Finally, our analysis offers a number of predictions that could be assessed if adequate data were available.

## References

- Acemoglu, D., G. Como, F. Fagnani, and A. Ozdaglar (2013). Opinion Fluctuations and Disagreement in Social Networks. *Mathematics of Operations Research* 38(1), 1–27.
- Acemoglu, D., A. Ozdaglar, and A. ParandehGheibi (2010). Spread of (Mis)Information in Social Networks. *Games and Economic Behavior* 70(2), 194–227.
- Akerlof, G. A. and R. Kranton (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton, NJ: Princeton University Press.
- Akerlof, G. A. and R. E. Kranton (2000). Economics and Identity. *Quarterly Journal of Economics* 115(3), 715–753.
- Alba, R. and V. Nee (1997). Rethinking Assimilation Theory for a New Era of Immigration. *International Migration Review* 31(4), 826–874.
- Algan, Y. and P. Cahuc (2006). Job Protection: The Macho Hypothesis. *Oxford Review of Economic Policy* 22(3), 390–410.
- Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006). Who’s Who in Networks. Wanted: The Key Player. *Econometrica* 74(5), 1403–1417.
- Böchel, B., T. Hellmann, and S. Klößner (2015). Opinion Dynamics and Wisdom under Conformity. *Journal of Economic Dynamics and Control* 52, 240–257.
- Bekkers, R. and P. Wiepking (2010). A Literature Review of Empirical Studies of Philanthropy: Eight Mechanisms that Drive Charitable Giving. *Nonprofit and Voluntary Sector Quarterly* 40(5), 924–973.
- Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou (2008). Are Muslim Immigrants Different in Terms of Cultural Integration? *Journal of the European Economic Association* 6(2-3), 445–456.
- Bisin, A. and T. Verdier (2000). “Beyond the Melting Pot”: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious traits. *The Quarterly Journal of Economics* 115(3), 955–988.
- Bisin, A. and T. Verdier (2001). The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory* 97(2), 298–319.
- Bisin, A. and T. Verdier (2010). The Economics of Cultural Transmission and Socialization. In J. Benhabib, A. Bisin, and M. Jackson (Eds.), *Handbook of Social Economics*. North-Holland.
- Borjas, G. J. (1985). Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants. *Journal of Labor Economics* 3(4), pp. 463–489.
- Breton, R. (1964). Institutional Completeness of Ethnic Communities and the Personal Relations of Immigrants. *American Journal of Sociology* 70(2), 193–205.
- Brubaker, R. (2001). The Return of Assimilation? Changing Perspectives on Immigration and its Sequels in France, Germany, and the United States. *Ethnic and Racial Studies* 24(4), 531–548.

- Büchel, B., T. Hellmann, and M. M. Pichler (2014). The Dynamics of Continuous Cultural traits in Social Networks. *Journal of Economic Theory* 154, 274–309.
- Carvalho, J.-P. and M. Koyama (2011). *Development and Religious Polarization: the Emergence of Reform and Ultra-Orthodox Judaism*. Department of Economics, University of Oxford.
- Cavalli-Sforza, L. L. and M. W. Feldman (1973). Cultural versus Biological Inheritance: Phenotypic Transmission from Parents to Children.(A Theory of the Effect of Parental Phenotypes on Children’s Phenotypes). *American Journal of Human Genetics* 25(6), 618.
- Ceylan, R. (2010). *Die Prediger des Islam Imame–Wer sie sind und was sie wirklich wollen* . Freiburg: Herder.
- Chatterjee, S. and E. Seneta (1977). Towards Consensus: Some Convergence Theorems on Repeated Averaging. *Journal of Applied Probability*, 89–97.
- Constant, A., L. Gataullina, K. Zimmermann, and L. Zimmermann (2006, September). Clash of Cultures: Muslims and Christians in the Ethnosizing Process. Discussion Paper 2350, IZA.
- DeMarzo, P., D. Vayanos, and J. Zwiebel (2003). Persuasion Bias, Social Influence, and Unidimensional Opinions. *The Quarterly Journal of Economics* 118(3), 909–968.
- Fernández, R. and A. Fogli (2006). Fertility: The Role of Culture and Family Experience. *Journal of the European Economic Association* 4(2-3), 552–561.
- Fernández, R., A. Fogli, and C. Olivetti (2004). Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *The Quarterly Journal of Economics* 119(4), 1249–1299.
- Fortin, N. M. (2005). Gender Role Attitudes and the Labor-Market Outcomes of Women across OECD Countries. *Oxford Review of Economic Policy* 21(3), 416–438.
- Fudenberg, D. and J. Tirole (1991). *Game Theory*. The MIT Press.
- Geaves, R. (2008). Drawing on the Past to Transform the Present: Contemporary Challenges for Training and Preparing British Imams. *Journal of Muslim Minority Affairs* 28(1), 99–112.
- Glazer, N. and D. P. Moynihan (1963). *Beyond the Melting Pot: The Negroes, Puerto Ricans, Jews, Italians and Irish of New York City*. The MIT Press.
- Guiso, L., P. Sapienza, and L. Zingales (2003). People’s Opium? Religion and Economic Attitudes. *Journal of Monetary Economics* 50(1), 225–282.
- Halm, D., M. Sauer, J. Schmidt, and A. Sticks (2012). Islamisches Gemeindeleben in Deutschland. Research report, Bundesamt für Migration und Flüchtlinge and Stiftung Zentrum für Türkeistudien und Integrationsforschung.
- Haug, S. (2008). *Sprachliche Integration von Migranten in Deutschland*. Number 14.
- Hauk, E. and H. Mueller (2015). Cultural Leaders and the Clash of Civilizations. *Journal of Conflict Resolution* 59, 367–400.

- Herberg, W. (1983). *Protestant–Catholic–Jew: An Essay in American Religious Sociology*. University of Chicago Press.
- Hofstede, G. (1984). *Culture's Consequences: International Differences in Work-Related Values*, Volume 5. Sage.
- Jackson, M. (2008). *Social and Economic Networks*. Princeton University Press.
- Judd, K. L. (1998). *Numerical Methods in Economics*. The MIT Press.
- Kantarevic, J. (2004, May). Interethnic Marriages and Economic Assimilation of Immigrants. Discussion Paper 1142, IZA.
- Katz, J. (1980). *From Prejudice to Destruction: Anti-Semitism, 1700-1933*. Harvard University Press.
- Koopmans, R. (2010). Trade-offs between Equality and Difference: Immigrant Integration, Multiculturalism and the Welfare State in Cross-National Perspective. *Journal of Ethnic and Migration Studies* 36(1), 1–26.
- Krueger, A. B. (2008). *What Makes a Terrorist: Economics and the Roots of Terrorism*. Princeton University Press.
- Kuran, T. and W. Sandholm (2008). Cultural Integration and Its Discontents. *Review of Economic Studies* 75(1), 201–228.
- Lorenz, J. (2005). A Stabilization Theorem for Dynamics of Continuous Opinions. *Physica A: Statistical Mechanics and its Applications* 355(1), 217–223.
- Lorenz, J. (2006). Convergence of Products of Stochastic Matrices with Positive Diagonals and the Opinion Dynamics Background. *Positive Systems* 341, 209–216.
- Mayer, E. (1979). *From Suburb to Shtetl: The Jews of Boro Park*. Transaction Publishers.
- Munshi, K. and N. Wilson (2008). Identity, Parochial Institutions, and Occupational Choice: Linking the Past to the Present in the American Midwest. Working Paper 13717, National Bureau of Economic Research.
- Panebianco, F. (2014). Socialization Networks and the Transmission of Interethnic Attitudes. *Journal of Economic Theory* 150, 583–610.
- Portes, A. and J. Sensenbrenner (1993). Embeddedness and Immigration: Notes on the Social Determinants of Economic Action. *American Journal of Sociology* 98(6), 1320–1350.
- Predelli, L. N. (2004). Interpreting Gender in Islam A Case Study of Immigrant Muslim Women in Oslo, Norway. *Gender & Society* 18(4), 473–493.
- Schalk-Soekar, S. R., F. J. van de Vijver, and M. Hoogsteder (2004). Attitudes toward Multiculturalism of immigrants and Majority Members in the Netherlands. *International Journal of Intercultural Relations* 28(6), 533–550.
- Subervi-Velez, F. A. (1986). The Mass Media and Ethnic Assimilation and Pluralism A Review and Research Proposal with Special Focus on Hispanics. *Communication Research* 13(1), 71–96.
- Vaughan, D. (2013). Nurture vs. Nurture: Endogenous Parental and Peer Effects and the Transmission of Culture. Working Paper 2013-04, Banco de México.

Verdier, T. and Y. Zenou (2015). The Role of Cultural Leaders in the Transmission of Preferences. *Economics Letters* 136, 158–161.

Verdier, T. and Y. Zenou (2016). Cultural Leaders and the Dynamics of Assimilation. Discussion Paper DP11174, CEPR.

Yasar, A. (2012). *Die DITIB zwischen der Türkei und Deutschland: Untersuchungen zur Türkisch-Islamischen Union der Anstalt für Religion eV*. Ergon.

Zhou, M. and G. Cai (2002). Chinese Language Media in the United States: Immigration and Assimilation in American Life. *Qualitative Sociology* 25(3), 419–441.

## A PROOFS

### Proof of Proposition 1 – Steady State in the Benchmark Case without Leader

Verify by inspection that  $\bar{p}_i = 0 \forall i$  is a solution to the steady state condition in equation (4).

To establish uniqueness and convergence, rewrite equation (3) in matrix form as follows:

$$p^{t+1} = \left[ \mathbf{I} - \hat{\mathbf{G}}(p^t, \alpha) \right] \mathbf{\Gamma} p^t,$$

where  $p^t$  is the vector of identities.  $\mathbf{I}$  is the identity matrix and we define  $\hat{\mathbf{G}}(p^t, \alpha)$  to be the  $n \times n$  matrix with diagonal elements  $\hat{\mathbf{G}}_{ii}(p^t) = \hat{g}(p_i^t; \alpha_i)$  and zeros elsewhere.  $\mathbf{\Gamma}$  is the group member influence matrix defined by  $\mathbf{\Gamma}_{ij} = \gamma_{ij}$ . Solving backwards yields:<sup>45</sup>

$$p^t = \prod_{s=0}^{t-1} \left[ \mathbf{I} - \hat{\mathbf{G}}(p^s, \alpha) \right] \cdot \mathbf{\Gamma}^t p^0.$$

Long run behavior of  $p^t$  as  $t \rightarrow \infty$  depends on the behavior of the two factors. We consider these separately. First  $\lim_{t \rightarrow \infty} \prod_{s=0}^{t-1} \left[ \mathbf{I} - \hat{\mathbf{G}}(p^s, \alpha) \right]$  is a diagonal matrix with element  $(i, i)$  given by  $\lim_{t \rightarrow \infty} \prod_{s=0}^{t-1} [1 - \hat{g}(p_i^s; \alpha_i)] = 0$  so that the matrix converges to a zero matrix. Second,  $\mathbf{\Gamma}$  is row stochastic with limit  $\lim_{t \rightarrow \infty} \mathbf{\Gamma}^t$  bounded above. Thus:

$$\lim_{t \rightarrow \infty} \prod_{s=0}^{t-1} \left[ \mathbf{I} - \hat{\mathbf{G}}(p^s, \alpha) \right] \mathbf{\Gamma}^t p^0 = 0$$

as required.

Note that alternatively our setting without leader can be rewritten as a repeated averaging process of  $n + 1$  players in which there is one column that contains strictly positive entries. The convergence result then follows from Theorem 4 of [Chatterjee and Seneta \(1977\)](#). That paper does not provide a characterization of the steady-state outcome.

### Proof of Proposition 2 – Speed of Convergence

We first prove the following result concerning the ranking of identities in subsequent time periods. Proposition 2 then follows immediately.

**Lemma 15.** *Consider a symmetric setting with  $k$  group members that all assign weight  $\gamma \in [\frac{1}{k}, 1]$  to their own identity and equal weight  $(1 - \gamma) / (k - 1)$  to each of the  $k - 1$  other members.*

<sup>45</sup>For this proof a superscript  $t$  applied to a matrix such as  $\mathbf{\Gamma}$  denotes  $\mathbf{\Gamma}$  taken to the  $t$ th power while a superscript on a vector such as  $p$  refers to the vector  $p$  in period  $t$ .

Then, for any pair of group members  $i, j$  such that  $\alpha_i \geq \alpha_j$ , if  $p_i^t \leq p_j^t$  this implies  $p_i^{t+1} \leq p_j^{t+1}$ . If the inequality on  $\alpha$  or on  $p^t$  or on both is strict, then  $p_i^{t+1} < p_j^{t+1}$ .

We show that the condition  $\gamma \in [\frac{1}{k}, 1]$  is sufficient for the result.

Take group members  $i, j$  such that  $\alpha_i \geq \alpha_j$  and let  $p_i^t \leq p_j^t$ . These assumptions imply  $\hat{g}(p_i^t; \alpha_i) \geq \hat{g}(p_j^t; \alpha_j)$ . Now, the difference between next period identities  $p_i^{t+1} - p_j^{t+1}$  is given by

$$\begin{aligned}
& p_i^{t+1} - p_j^{t+1} \\
&= [1 - \hat{g}(p_i^t; \alpha_i)] \left\{ \gamma p_i^t + (1 - \gamma)/(k - 1) \sum_{l \neq i} p_l^t \right\} \\
&\quad - [1 - \hat{g}(p_j^t; \alpha_j)] \left\{ \gamma p_j^t + (1 - \gamma)/(k - 1) \sum_{l \neq j} p_l^t \right\} \\
&= [\hat{g}(p_j^t; \alpha_j) - \hat{g}(p_i^t; \alpha_i)] \left\{ (1 - \gamma)/(k - 1) \sum_{l \neq i, j} p_l^t \right\} \\
&\quad + \left\{ [1 - \hat{g}(p_i^t; \alpha_i)] \left[ \gamma p_i^t + \frac{1 - \gamma}{k - 1} p_j^t \right] - [1 - \hat{g}(p_j^t; \alpha_j)] \left[ \gamma p_j^t + \frac{1 - \gamma}{k - 1} p_i^t \right] \right\}
\end{aligned}$$

The first part of this sum is non-positive given  $\hat{g}(p_j^t; \alpha_j) - \hat{g}(p_i^t; \alpha_i) \leq 0$ . Furthermore, the second part consists of the difference between two weighted averages of  $p_i^t$  and  $p_j^t$ . If  $\gamma \geq 1/k$ , then  $\gamma \geq \frac{1-\gamma}{k-1}$  so that the first of these averages is weakly smaller than the second. In addition  $\hat{g}(p_j^t; \alpha_j) - \hat{g}(p_i^t; \alpha_i) \leq 0$  implies  $[1 - \hat{g}(p_i^t; \alpha_i)] \leq [1 - \hat{g}(p_j^t; \alpha_j)]$  which ensures that also the second part of the sum is negative. The sum of two non-positive numbers is non-positive and thus overall  $p_i^{t+1} - p_j^{t+1} \leq 0$  as required.

The result for strict inequality follows from recognizing that  $\hat{g}(p_j^t; \alpha_j) - \hat{g}(p_i^t; \alpha_i) < 0$  if either  $\alpha_i > \alpha_j$  or  $p_i^t < p_j^t$  or both. In that case, both parts of the sum are strictly negative and thus  $p_i^{t+1} - p_j^{t+1} < 0$ .

Lemma 15 shows that if group members start with the same identity but have different ability, then under the conditions of Proposition 2 the high ability group member will pull ahead. Furthermore, if the more integrated player has at least the same ability as the other, the ranking of two group members' identities is maintained across two subsequent time periods. It follows that the rankings are also maintained for periods more than one step apart and we then arrive at the result stated in Proposition 2.

## Proof of Proposition 3 - Steady State with Leader

### Existence and Characterization as Convex Combination

Rewrite the updating process of the vector of group member identities into a new identity vector as  $p' = \Phi(p)$ . A steady state identity vector is then a fixed point of  $\Phi(\cdot)$ .

The domain and co-domain of  $\Phi(\cdot)$  are the closed interval  $[0, p^{\max}]^n$ . Continuity of  $\Phi(\cdot)$  follows from its definition in equation (6). Therefore by Brouwer's Fixed Point Theorem,  $\Phi(\cdot)$  has a fixed point and there exists a steady state.

To show  $p_i \in (0, p_L) \forall i$  we consider each boundary separately and proceed by contradiction. The result then follows.

First, suppose  $\tilde{p}$  is a fixed point of  $\Phi(p)$  and there exists a player such that  $\tilde{p}_i = 0$ . Then

$$\Phi_i(\tilde{p}) = [1 - \hat{g}(\tilde{p}_i; \alpha_i)] \left[ \lambda_i p_L + (1 - \lambda_i) \sum_{j=1}^n \gamma_{ij} p_j \right]$$

which is strictly greater than zero because  $\hat{g}(\tilde{p}_i; \alpha_i) < 1$ ,  $\lambda_i > 0$  and  $p_L > 0$ .  $\tilde{p}$  thus is not a fixed point, delivering the contradiction.

Second, suppose  $\tilde{p}$  is a fixed point of  $\Phi(p)$  and there is a player with  $\tilde{p}_i \geq p_L$ . Label the player with the highest  $\tilde{p}_i$  as  $i_{\max}$ . Then  $\tilde{p}_{i_{\max}} \geq p_L$ . Again

$$\Phi_{i_{\max}}(\tilde{p}) = [1 - \hat{g}(\tilde{p}_{i_{\max}}; \alpha_{i_{\max}})] \left[ \lambda_i p_L + (1 - \lambda_i) \sum_{j=1}^n \gamma_{ij} p_j \right]$$

which is strictly less than  $\tilde{p}_{i_{\max}}$  as  $\hat{g}(\tilde{p}_i; \alpha_i) > 0$ . This implies  $\tilde{p}$  is not a fixed point, delivering the contradiction.

### Uniqueness and Convergence

We show that the updating process with leader described by equation (6) is a contraction under a suitable norm and then appeal to Blackwell's contraction mapping theorem to show uniqueness and convergence.

Let  $\Phi(p)$  be the one period updating process of group member identity vector  $p$  yielding next period vector  $p'$ . Then for  $\Phi(p)$  to be a contraction we need to show that for every two

$n$ -dimensional identity vectors  $p \neq q$  and for some norm  $\|\cdot\|$  and scalar  $c < 1$ :

$$\|\Phi(p) - \Phi(q)\| \leq c \|p - q\|.$$

Equation (6) and assumptions on  $\hat{g}(\cdot)$  imply that  $\Phi(p)$  is continuous and differentiable everywhere. We can then prove that  $\Phi(\cdot)$  is a contraction via the Jacobian  $J(\Phi(p))$  if there exists a matrix norm  $\|\cdot\|$  of  $J$  and scalar  $c < 1$  such that for every  $p$ :<sup>46</sup>

$$\|J(\Phi(p))\| \leq c.$$

We use is the matrix norm induced by the  $\infty$ -vector norm defined for matrix  $A$  as:

$$\|A\|_{\infty} = \max_i \left[ \sum_j A_{ij} \right].$$

$\Phi(p)$  is therefore a contraction if:

$$\max_i \left[ \sum_j J_{ij}(\Phi(p)) \right] < 1.$$

Computing the elements of the Jacobian

$$\begin{aligned} \frac{\partial \Phi_i(p)}{\partial p_i} &= [1 - \hat{g}(p_i; \alpha_i)] (1 - \lambda_i) \gamma_{ii} + \left\{ \lambda_i p_L + (1 - \lambda_i) \sum_j (\gamma_{ij} p_j) \right\} \left[ -\alpha_i \frac{\partial \hat{g}(p_i; \alpha_i)}{\partial p_i} \right] \\ \frac{\partial \Phi_i(p)}{\partial p_j} &= [1 - \hat{g}(p_i; \alpha_i)] (1 - \lambda_i) \sum_{j \neq i} \gamma_{ij}, \end{aligned}$$

we derive the following sufficient condition:

$$\begin{aligned} &\|\mathbf{J}(\Phi(p))\| \\ &= \max_i \left[ [1 - \hat{g}(p_i; \alpha_i)] (1 - \lambda_i) + \left\{ \lambda_i p_L + (1 - \lambda_i) \sum_j (\gamma_{ij} p_j) \right\} \left[ -\frac{\partial \hat{g}(p_i; \alpha_i)}{\partial p_i} \right] \right] \\ &\leq \max_i \left[ [1 - \hat{g}(p_i; \alpha_i)] + p^{\max} \left[ -\frac{\partial \hat{g}(p_i; \alpha_i)}{\partial p_i} \right] \right] \\ &< 1. \end{aligned}$$

---

<sup>46</sup>See Judd (1998, Theorem 5.4.1) for the approach adopted here.

The line following the first  $\leq$  provides an upper bound for all network structures, all initial values of  $p$  and all  $\lambda_i$  (in particular  $\lambda_i \rightarrow 0$ ). This condition further simplifies to:

$$p^{\max} < \frac{\hat{g}(p_i; \alpha_i)}{\left| \frac{\partial \hat{g}(p_i; \alpha_i)}{\partial p_i} \right|} \quad \forall i \in N \text{ and } \forall p_i \in [0, p^{\max}].$$

which delivers Assumption 1. It then follows from the contraction mapping theorem that  $\Phi(\cdot)$  has a unique steady state and the system converges globally to the steady state.

### Proof of Proposition 4 - Network Structure and Steady State

First note that we can summarize the updating process from period  $t$  to  $t + 1$  of the model with leader in the following matrix form

$$p_+^{t+1} = \mathbf{X}(p_+^t) p_+^t$$

where  $p_+^t$  is a  $2+n$  vector of identities constructed by adding the host society and the leader to  $p^t$  and the updating matrix is given by:

$$\begin{aligned} & \mathbf{X}(H(t)) \\ &= \left( \begin{array}{cc|ccc} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \hline \hat{g}(p_1^t) & (1 - \hat{g}(p_1^t))\lambda_1 & (1 - \hat{g}(p_1^t))(1 - \lambda_1)\gamma_{1,1} & \cdots & (1 - \hat{g}(p_1^t))(1 - \lambda_1)\gamma_{1,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{g}(p_n^t) & (1 - \hat{g}(p_n^t))\lambda_n & (1 - \hat{g}(p_n^t))(1 - \lambda_n)\gamma_{n,1} & \cdots & (1 - \hat{g}(p_n^t))(1 - \lambda_n)\gamma_{n,n} \end{array} \right) \\ &= \left( \begin{array}{cc|ccc} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \hline \hat{\mathbf{G}} \mathbf{1} & [\mathbf{I} - \hat{\mathbf{G}}(\bar{p})] \mathbf{\Lambda} \mathbf{1} & [\mathbf{I} - \hat{\mathbf{G}}(\bar{p})] & [\mathbf{I} - \mathbf{\Lambda}] \mathbf{\Gamma} & \end{array} \right) \\ &\equiv \left( \begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{C}(p^t) & \mathbf{D}(p^t) \end{array} \right) \end{aligned}$$

where  $\mathbf{I}$  is a  $2 \times 2$  identity matrix,  $\mathbf{0}$  a  $n \times 2$  zero matrix and  $\mathbf{\Lambda}$  and  $\hat{\mathbf{G}}(\bar{p})$  are diagonal matrices with entries  $\lambda$  and  $\hat{g}(p)$  on the diagonal, respectively.  $\mathbf{\Gamma}$  presents the row stochastic influence matrix amongst the group members.

$\mathbf{C}(p^t)$  is then defined as a  $n \times 2$  matrix describing the interaction of group members with

the host society and the leader, and  $\mathbf{D}(p^t)$  as a  $n \times n$  matrix that reflects interaction of group members amongst each other.

Now, consider  $\mathbf{X}(t)$  which provides the  $t$ -th iteration with multiplication from the left of  $\mathbf{X}(p^s)$  for a sequence of  $p^s$  from  $s = 0$  to  $t$ . This can be partitioned into

$$\begin{aligned}\mathbf{X}(t) &= \prod_{s=0}^t (\mathbf{X}(p^s)) \\ &= \left( \begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{C}(t) & \mathbf{D}(t) \end{array} \right)\end{aligned}$$

where

$$\begin{aligned}\mathbf{C}(t) &= \mathbf{D}(p^t)\mathbf{C}(t-1) + \mathbf{C}(p^t) \\ \mathbf{D}(t) &= \prod_{s=0}^t (\mathbf{D}(p^s))\end{aligned}$$

We consider the limits of  $\mathbf{C}(t)$  and  $\mathbf{D}(t)$  as  $t \rightarrow \infty$ . Start with  $\mathbf{D}(t)$ . The sum of each row of  $\mathbf{D}(p^s)$  is less than one for strictly positive  $\hat{g}(p_i^t; \alpha_i)$  and  $\lambda_i \in (0, 1)$ , which implies convergence to zero in the sense that

$$\lim_{t \rightarrow \infty} \mathbf{D}(t) = \lim_{t \rightarrow \infty} \prod_{s=0}^t (\mathbf{D}(p^s)) = \mathbf{0}$$

where here  $\mathbf{0}$  is an  $n \times n$  zero matrix.

Next consider  $\lim_{t \rightarrow \infty} \mathbf{C}(t)$ . Given  $\lim_{t \rightarrow \infty} \mathbf{D}(t) = \mathbf{0}$ , we have  $\lim_{t \rightarrow \infty} \mathbf{C}(t) \equiv \bar{\mathbf{C}}$  finite and can characterize it inductively as follows:

$$\begin{aligned}\bar{\mathbf{C}} &= \mathbf{D}(\bar{p})\bar{\mathbf{C}} + \mathbf{C}(\bar{p}) \\ &= [\mathbf{I} - \mathbf{D}(\bar{p})]^{-1} \mathbf{C}(\bar{p})\end{aligned}$$

Thus, we have

$$\bar{\mathbf{X}} = \lim_{t \rightarrow \infty} \mathbf{X}(t) = \left( \begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \bar{\mathbf{C}} & \mathbf{0} \end{array} \right)$$

and  $\bar{p}_+ = \bar{\mathbf{X}}p_+^0$ . The  $\mathbf{0}$  in the bottom right block of  $\bar{\mathbf{X}}$  implies that  $\bar{p}$ , that is, the vector of long-run entries of  $\bar{p}_+$  corresponding to the group members, are a convex combination of

$p_S$  and  $p_L$  only, so that:

$$\begin{aligned}\bar{p} &= \bar{\mathbf{C}} \begin{pmatrix} p_S \\ p_L \end{pmatrix} \\ &= [\mathbf{I} - \mathbf{D}(\bar{p})]^{-1} \mathbf{C}(\bar{p}) \begin{pmatrix} 0 \\ p_L \end{pmatrix} \\ &= [\mathbf{I} - \mathbf{D}(\bar{p})]^{-1} \mathbf{C}(\bar{p})_{[,2]} p_L\end{aligned}$$

where  $\mathbf{C}(\bar{p})_{[,2]}$  refers to the second column of the matrix  $\mathbf{C}(\bar{p})$ .

Substituting the definitions of matrices  $\mathbf{D}(p)$  and  $\mathbf{C}(p)$  from above yields

$$\bar{p} = \left\{ \mathbf{I} - (\mathbf{I} - \mathbf{\Lambda}) \left[ \mathbf{I} - \hat{\mathbf{G}}(\bar{p}) \right] \mathbf{\Gamma} \right\}^{-1} [\mathbf{I} - \hat{\mathbf{G}}(\bar{p})] \mathbf{\Lambda} \mathbf{1} p_L$$

as required.

### Proof of Proposition 6: Identity Spread for Different Types

The results follow from the implicit function theorem applied to the system of equations describing the steady state (equations (9) and (10)). We define

$$F_A(\bar{p}_A, \bar{p}_B; \gamma, p_L) \equiv [1 - \hat{g}(\bar{p}_A; \alpha_A)] \{ \lambda p_L + (1 - \lambda) [\gamma \bar{p}_A + (1 - \gamma) \bar{p}_B] \} - \bar{p}_A = 0$$

$$F_B(\bar{p}_A, \bar{p}_B; \gamma, p_L) \equiv [1 - \hat{g}(\bar{p}_B; \alpha_B)] \{ \lambda p_L + (1 - \lambda) [\gamma \bar{p}_B + (1 - \gamma) \bar{p}_A] \} - \bar{p}_B = 0.$$

By the implicit function theorem the effect of a change in the parameter  $\gamma$  on  $i$ 's identity is given by:

$$\frac{\partial p_i}{\partial \gamma} = \frac{-|J_i(\gamma)|}{|J|}$$

where  $J$  is the Jacobian of  $F$  with respect to endogenous variables and  $J_i(\theta)$  is the same matrix with column  $i$  replaced by the vector of partial derivatives with respect to parameter

$\theta$ . We compute  $|J|$  and  $|J_A(\gamma)|$  from the partial derivatives given by:

$$\frac{\partial F_i}{\partial \bar{p}_i} = [1 - \hat{g}(\bar{p}_i; \alpha_i)] (1 - \lambda)\gamma - \frac{\partial \hat{g}(\bar{p}_i; \alpha_i)}{\partial p_i} \{ \lambda p_L + (1 - \lambda) [\gamma \bar{p}_i + (1 - \gamma) \bar{p}_j] \} - 1 < 0 \quad (15)$$

$$\frac{\partial F_i}{\partial \bar{p}_j} = [1 - \hat{g}(\bar{p}_i; \alpha_i)] (1 - \lambda)(1 - \gamma) > 0 \quad (16)$$

$$\frac{\partial F_i}{\partial \gamma} = [1 - \hat{g}(\bar{p}_i; \alpha_i)] (1 - \lambda) (\bar{p}_i - \bar{p}_j). \quad (17)$$

Equation (15) can be signed using Assumption 1. The sign of equation (16) follows directly from  $\hat{g}(\bar{p}_i; \alpha_i) \in (0, 1)$ . Equation (17) adopts the sign of  $(\bar{p}_i - \bar{p}_j)$  which is negative for  $\frac{\partial F_A}{\partial \gamma}$ .

The determinant of the Jacobian  $|J|$  is then given by

$$|J| = \frac{\partial F_A}{\partial p_A} \frac{\partial F_B}{\partial p_B} - \frac{\partial F_A}{\partial p_B} \frac{\partial F_B}{\partial p_A} > 0.$$

The inequality follows by recognizing that Assumption 1 implies that  $|\frac{\partial F_i}{\partial p_i}| > |\frac{\partial F_i}{\partial p_j}|$  for both  $i = A$  and  $i = B$ .

1. *In the model with two-types, if  $\alpha_A > \alpha_B$  then  $p_A < p_B$  for all levels of inter-connection  $\gamma$  between types.*

As  $|J| > 0$ , the Jacobian is non-singular for all  $\bar{p}_A, \bar{p}_B \in [0, p^{\max}]$ . By the implicit function theorem the mapping of parameters to long run identities  $\bar{p}_A$  and  $\bar{p}_B$  is therefore continuously differentiable over the domain of  $\gamma$ . Differentiability implies continuity. Proceed by contradiction. Define  $\bar{p}_B \equiv \bar{p}_A + \Delta$  and suppose there exists  $\hat{\gamma}$  such that  $\bar{p}_B \leq \bar{p}_A$  and thus  $\Delta \leq 0$ . If  $\gamma = 1$ , then  $\Delta > 0$ . Given continuity of the implicit function it follows by the intermediate value theorem that there must exist  $\tilde{\gamma} \in [\hat{\gamma}, 1]$  such that  $\Delta = 0$  or equivalently  $\bar{p}_A = \bar{p}_B = \tilde{p}$ . By the definition of the steady state this implies:

$$\begin{aligned} [1 - \hat{g}(\tilde{p}; \alpha_A)] \{ \lambda p_L + (1 - \lambda) \tilde{p} \} &= [1 - \hat{g}(\tilde{p}; \alpha_B)] \{ \lambda p_L + (1 - \lambda) \tilde{p} \} \\ \Leftrightarrow \hat{g}(\tilde{p}; \alpha_A) &= \hat{g}(\tilde{p}; \alpha_B) \end{aligned}$$

where the final line contradicts the assumption that  $\alpha_A > \alpha_B$ .

2.  $|\bar{p}_A - \bar{p}_B|$  decreases in  $\gamma$ .

The determinant of the modified Jacobian  $J_A(\gamma)$  can be signed as follows:

$$\begin{aligned}
|J_A(\gamma)| &= \frac{\partial F_A}{\partial \gamma} \frac{\partial F_B}{\partial p_B} - \frac{\partial F_A}{\partial p_B} \frac{\partial F_B}{\partial \gamma} \\
&= [1 - \hat{g}(\bar{p}_A; \alpha_A)] (1 - \lambda) (\bar{p}_A - \bar{p}_B) \cdot \\
&\quad \left\{ [1 - \hat{g}(\bar{p}_B; \alpha_B)] (1 - \lambda) - \frac{\partial \hat{g}(\bar{p}_B; \alpha_B)}{\partial \bar{p}_B} \{ \lambda p_L + (1 - \lambda) [\gamma \bar{p}_B + (1 - \gamma) \bar{p}_A] \} - 1 \right\} \\
&> 0
\end{aligned}$$

The final inequality is due to the multiplication of two negative components: The first part has the sign of  $\bar{p}_A - \bar{p}_B$ , which is negative as  $\alpha_A > \alpha_B$ . The second part is negative by Assumption 1, applied here to  $\bar{p}_B$ .

Reassembling, we can now sign the effect of  $\gamma$  on  $\bar{p}_A$ :

$$\frac{\partial \bar{p}_A}{\partial \gamma} = \frac{-|J_A(\gamma)|}{|J|} < 0.$$

The proof that  $\frac{\partial \bar{p}_B}{\partial \gamma} = \frac{-|J_B(\gamma)|}{|J|} > 0$  works along the same lines and is omitted for brevity.

### Proof of Proposition 7: Leader Decision

We first establish that  $\bar{p}(p_L)$  is an increasing function in  $p_L$ . To see this recall equation (11)

$$\bar{p} = [1 - \hat{g}(\bar{p}; \alpha)] \{ \lambda p_L + (1 - \lambda) \bar{p} \}.$$

Differentiating with respect to  $p_L$  yields

$$\frac{\partial \bar{p}}{\partial p_L} = \frac{\lambda (1 - \hat{g}(\bar{p}; \alpha))}{1 + \frac{\partial \hat{g}(\bar{p}; \alpha)}{\partial \bar{p}} (\lambda p_L + (1 - \lambda) \bar{p}) - (1 - \lambda) (1 - \hat{g}(\bar{p}; \alpha))}$$

which is strictly positive under Assumption 1. This strictly increasing relationship between  $p_L$  and  $\bar{p}$  allows us to study the leader problem from the perspective of  $\bar{p}$ . Note also that even if  $p_L = p^{\max}$  the resulting  $\bar{p}$  is strictly below  $p^{\max}$ .

To simplify notation, we define the steady state income of group members as

$$\bar{w} \equiv \alpha \hat{f}(\bar{p}; \alpha).$$

The leader's payoff is then given by  $\pi(\bar{p}, \bar{w})$  and the first order condition is

$$\left( \frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{p}} + \frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}} \frac{\partial \bar{w}}{\partial \bar{p}} \right) \frac{\partial \bar{p}}{\partial p_L}.$$

We can have three different cases:

1. The leader sets  $p_L = p^{\max}$  if

$$\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{p}} + \frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}} \frac{\partial \bar{w}}{\partial \bar{p}} > 0 \quad \Leftrightarrow \quad \frac{\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{p}}}{\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}}} > -\alpha \frac{\partial \hat{f}(\bar{p}; \alpha)}{\partial \bar{p}}.$$

2. He sets  $p_L = 0$  if

$$\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{p}} + \frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}} \frac{\partial \bar{w}}{\partial \bar{p}} < 0 \quad \Leftrightarrow \quad \frac{\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{p}}}{\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}}} < -\alpha \frac{\partial \hat{f}(\bar{p}; \alpha)}{\partial \bar{p}}.$$

3. He sets  $p_L \in [0, p^{\max}]$  if

$$\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{p}} + \frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}} \frac{\partial \bar{w}}{\partial \bar{p}} = 0.$$

In case the first order condition equals zero for some  $\bar{p}$ , we have an interior global maximum if  $\frac{\partial^2 \hat{f}(\bar{p}; \alpha)}{\partial \bar{p}^2} < 0$  which is equivalent to the condition that  $f'''(\bar{H}) < 0$ . To see this, consider the second order equation:

$$\underbrace{\frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p}^2}}_{\leq 0} + 2 \underbrace{\frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p} \partial \bar{w}} \frac{\partial \bar{w}}{\partial \bar{p}}}_{\leq 0} + \frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}} \frac{\partial^2 \bar{w}}{\partial \bar{p}^2} + \underbrace{\frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{w}^2} \left( \frac{\partial \bar{w}}{\partial \bar{p}} \right)^2}_{\leq 0}.$$

The second order condition is negative whenever:

$$\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}} \frac{\partial^2 \bar{w}}{\partial \bar{p}^2} < 0,$$

where  $\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}} > 0$  and

$$\frac{\partial^2 \bar{w}}{\partial \bar{p}^2} = \alpha \frac{\partial^2 \hat{f}(\bar{p}, \alpha)}{\partial \bar{p}^2} = \alpha \left( \frac{c_1^2}{\alpha^2 f''(\bar{H})} \right) \left\{ 1 - f'(\bar{H}) \frac{f'''(\bar{H})}{[f''(\bar{H})]^2} \right\}.$$

Thus, a sufficient condition for an interior solution that maximizes leader's payoff is  $\frac{\partial^2 \hat{f}(\bar{p}; \alpha)}{\partial \bar{p}^2} < 0$  which is implied by  $f'''(\bar{H}) < 0$ . If this condition is violated the interior solution might

be a minimum and the leader will select an extreme solution at either  $p_L = p^{\max}$  or  $p_L = 0$ .

**Strategic Substitutes** If identity and earnings are strategic substitutes, then the analysis of parts (1) and (2) remains unchanged. Part (3) is affected, though. It is no longer sufficient that  $f'''(\bar{H}) < 0$ . However, the key results, of leaders choosing either between complete cultural differentiation, intermediate differentiation or full integration remains unchanged.

### Proof of Proposition 9: Comparative Statics $\alpha$

**Effect on  $p_L$**  In case of the interior solution being a maximum, we consider the effect of an increase of  $\alpha$  on  $p_L$ , or rather  $\bar{p}$ , as the leader can set directly the group member identity he prefers.

We set  $F \equiv \left( \frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{p}} + \frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}} \frac{\partial \bar{w}}{\partial \bar{p}} \right) = 0$ . From this we can calculate the effect of  $\alpha$  on  $\bar{p}$ ,

$$\frac{\partial \bar{p}}{\partial \alpha} = - \frac{\frac{\partial F}{\partial \alpha}}{\frac{\partial F}{\partial \bar{p}}}.$$

At an interior solution we have  $\frac{\partial F}{\partial \bar{p}} < 0$  and thus  $\text{sign} \left( \frac{\partial \bar{p}}{\partial \alpha} \right) = \text{sign} \left( \frac{\partial F}{\partial \alpha} \right)$ , where

$$\frac{\partial F}{\partial \alpha} = \underbrace{\left( \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p} \partial \bar{w}} + \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{w}^2} \frac{\partial \bar{w}}{\partial \bar{p}} \right)}_{>0} \frac{\partial \bar{w}}{\partial \alpha} + \underbrace{\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}}}_{>0} \frac{\partial^2 \bar{w}}{\partial \bar{p} \partial \alpha}.$$

Thus, the sign of the derivative depends on  $\frac{\partial \bar{w}}{\partial \alpha}$  and  $\frac{\partial^2 \bar{w}}{\partial \bar{p} \partial \alpha}$ . We know that:

$$\frac{\partial \bar{w}}{\partial \alpha} = f(\bar{H}) + \alpha f'(\bar{H}) \frac{\partial \bar{H}}{\partial \alpha} > 0.$$

Further,

$$\frac{\partial^2 \bar{w}}{\partial \bar{p} \partial \alpha} = f'(\bar{H}) \frac{\partial \bar{H}}{\partial \bar{p}} + \alpha f''(\bar{H}) \frac{\partial \bar{H}}{\partial \bar{p}} \frac{\partial \bar{H}}{\partial \alpha} + \alpha f'(\bar{H}) \frac{\partial^2 \bar{H}}{\partial \bar{p} \partial \alpha},$$

where

$$\frac{\partial \bar{H}}{\partial \bar{p}} = \frac{c_1}{\alpha f''(\bar{H})}, \quad \frac{\partial \bar{H}}{\partial \alpha} = - \frac{f'(\bar{H})}{\alpha f''(\bar{H})}, \quad \frac{\partial^2 \bar{H}}{\partial \bar{p} \partial \alpha} = - \frac{c_1}{\alpha^2 f''(\bar{H})} \left( 1 - \frac{f'(\bar{H}) f'''(\bar{H})}{[f''(\bar{H})]^2} \right).$$

Then,

$$\frac{\partial^2 \bar{w}}{\partial \bar{p} \partial \alpha} = -\frac{c_1}{\alpha} \frac{f'(\bar{H})}{f''(\bar{H})} \left( 1 - \frac{f'(\bar{H}) f'''(\bar{H})}{[f''(\bar{H})]^2} \right) > 0,$$

which implies that  $\text{sign} \left( \frac{\partial F}{\partial \alpha} \right) = \text{sign} \left( \frac{\partial \bar{p}}{\partial \alpha} \right) > 0$  as we have assumed that  $\left( 1 - \frac{f'(\bar{H}) f'''(\bar{H})}{[f''(\bar{H})]^2} \right) > 0$  in order to ensure that the interior solution is a maximum. Thus, an increase in  $\alpha$  leads to a higher  $p_L$ .

**Effect on  $\bar{H}$**  The effect of an increase of  $\alpha$  on  $\bar{H}$  is given by  $\frac{d\bar{H}}{d\alpha} = \frac{1}{\alpha f''(\bar{H})} \left( c_1 \frac{d\bar{p}}{d\alpha} - f'(\bar{H}) \right)$ , which implies that  $\text{sign} \left( \frac{d\bar{H}}{d\alpha} \right) = -\text{sign} \left( c_1 \frac{d\bar{p}}{d\alpha} - f'(\bar{H}) \right)$ , with

$$\begin{aligned} c_1 \frac{d\bar{p}}{d\alpha} - f'(\bar{H}) &= -f'(\bar{H}) \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p}^2} + \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p} \partial \bar{w}} \left( -c_1 f(\bar{H}) - f'(\bar{H}) \frac{\partial \bar{w}}{\partial \bar{p}} \right) \\ &\quad + \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{w}^2} \frac{\partial \bar{w}}{\partial \bar{p}} (-c_1 f(\bar{H})). \end{aligned} \quad (18)$$

The sign of equation (18) is ambiguous without further restrictions.

**Effect on  $\bar{w}$**  We show that an increase in  $\alpha$  leads to higher income. The effect of  $\alpha$  on  $\bar{w}$  is given by:

$$\frac{\partial \bar{w}}{\partial \alpha} = f(\bar{H}) + \frac{f'(\bar{H})}{f''(\bar{H})} \left( -c_1 \frac{\partial F}{\partial \alpha} - f'(\bar{H}) \right) > 0$$

which is equivalent to

$$\frac{f(\bar{H}) f''(\bar{H})}{f'(\bar{H})} \frac{\partial F}{\partial \bar{p}} > c_1 \frac{\partial F}{\partial \alpha} + f' \frac{\partial F}{\partial \bar{p}}. \quad (19)$$

We know from equation (18) that the term on the right can be rewritten as

$$c_1 \frac{\partial F}{\partial \alpha} + f' \frac{\partial F}{\partial \bar{p}} = f'(\bar{H}) \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p}^2} + \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p} \partial \bar{w}} \left( c_1 f(\bar{H}) + f'(\bar{H}) \frac{\partial \bar{w}}{\partial \bar{p}} \right) + \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{w}^2} \frac{\partial \bar{w}}{\partial \bar{p}} (c_1 f(\bar{H})).$$

Equation (19) can then be rewritten as

$$\begin{aligned} \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p}^2} \left( \frac{f(\bar{H}) f''(\bar{H})}{f'(\bar{H})} - f'(\bar{H}) \right) &+ \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p} \partial \bar{w}} \left( 2 \frac{f(\bar{H}) f''(\bar{H})}{f'(\bar{H})} \frac{\partial \bar{w}}{\partial \bar{p}} - c_1 f'(\bar{H}) - f'(\bar{H}) \frac{\partial \bar{w}}{\partial \bar{p}} \right) \\ &+ \frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}} \frac{\partial^2 \bar{w}}{\partial \bar{p}^2} \frac{f(\bar{H}) f''(\bar{H})}{f'(\bar{H})} + \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{w}^2} \frac{\partial \bar{w}}{\partial \bar{p}} \left( \frac{f(\bar{H}) f''(\bar{H})}{f'(\bar{H})} \frac{\partial \bar{w}}{\partial \bar{p}} - c_1 f(\bar{H}) \right) > 0. \end{aligned}$$

Taking into account that  $\frac{\partial \bar{w}}{\partial \bar{p}} = \frac{c_1 f'(\bar{H})}{f''(\bar{H})} < 0$  and the signs of the derivatives which have been established before, we find that  $\frac{\partial \bar{w}}{\partial \alpha} > 0$ .

### Proof of Corollary 9.1: Increase in Skill Investment

We show that an increase in  $\alpha$  leads to an increase in skill investment for the business leader.

By equation (18):

$$\begin{aligned} \text{sign}\left(\frac{d\bar{H}}{d\alpha}\right) &= \text{sign}\left(-f'(\bar{H})\frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p}^2} + \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p} \partial \bar{w}} \left(-c_1 f(\bar{H}) - f'(\bar{H})\frac{\partial \bar{w}}{\partial \bar{p}}\right)\right. \\ &\quad \left.+ \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{w}^2} \frac{\partial \bar{w}}{\partial \bar{p}} (-c_1 f(\bar{H}))\right). \end{aligned}$$

The derivatives are given by:

$$\begin{aligned} \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p}^2} &= -\beta(1-\beta)\bar{p}^{\beta-2}\bar{w}^{1-\beta} \\ \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p} \partial \bar{w}} &= \beta(1-\beta)\bar{p}^{\beta-1}\bar{w}^{-\beta} \\ \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{w}^2} &= -\beta(1-\beta)\bar{p}^{\beta}\bar{w}^{-\beta-1}. \end{aligned}$$

We can then write the statement of the corollary as

$$f'(\bar{H})\bar{p}^{-1}\bar{w} + \left(-c_1 f(\bar{H}) - f'(\bar{H})\frac{\partial \bar{w}}{\partial \bar{p}}\right) + \bar{p}\bar{w}^{-1}c_1 f(\bar{H})\frac{\partial \bar{w}}{\partial \bar{p}} > 0$$

which after some algebra and substituting for the definition of  $w$  simplifies to:

$$f'(\bar{H})\frac{\alpha}{\bar{p}} = \frac{c_0}{\bar{p}} + c_1 > c_1.$$

This establishes the result.

### Proof of Proposition 10: Effect of a Lump Sum Transfer

This proof follows along the lines of the Proof of Proposition 9, with the only change being the lump sum transfer added to income from skills and thus  $\bar{w} = \alpha f(H^*) + z$ . The sign of a

change in the transfer  $z$  is given by  $sign\left(\frac{\partial \bar{p}}{\partial z}\right) = sign\left(\frac{\partial F}{\partial z}\right)$ , where

$$\frac{\partial F}{\partial z} = \underbrace{\left(\frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p} \partial \bar{w}} + \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{w}^2} \frac{\partial \bar{w}}{\partial \bar{p}}\right)}_{\geq 0} \frac{\partial \bar{w}}{\partial z} + \underbrace{\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}}}_{> 0} \frac{\partial^2 \bar{w}}{\partial \bar{p} \partial z}$$

We know that

$$\frac{\partial \bar{w}}{\partial z} = \frac{\partial(\alpha \hat{f}(\bar{p}; \alpha) + z)}{\partial z} = 1, \quad \frac{\partial^2 \bar{w}}{\partial \bar{p} \partial z} = 0$$

as  $\frac{\partial \bar{w}}{\partial \bar{p}}$  is independent of  $z$ . This implies that  $\frac{\partial F}{\partial z} \geq 0$  and thus  $\frac{\partial \bar{p}}{\partial z} \geq 0$ . Then

$$\frac{\partial H^*}{\partial z} = \frac{c_1}{\alpha f''(H^*)} \frac{\partial \bar{p}}{\partial z} \leq 0.$$

The effect of the transfer on wages is then given by:

$$\frac{\partial \bar{w}}{\partial z} = \alpha f'(H^*) \frac{\partial H^*}{\partial z} + 1 = \frac{\partial \bar{p}}{\partial z} \frac{\partial \bar{w}}{\partial \bar{p}} + 1 = -\frac{\frac{\partial F}{\partial z}}{\frac{\partial F}{\partial \bar{p}}} \frac{\partial \bar{w}}{\partial \bar{p}} + 1.$$

Wages are positive if

$$1 > \frac{\frac{\partial F}{\partial z}}{\frac{\partial F}{\partial \bar{p}}} \frac{\partial \bar{w}}{\partial \bar{p}} \quad \Leftrightarrow \quad \frac{\partial F}{\partial z} \frac{\partial \bar{w}}{\partial \bar{p}} > \frac{\partial F}{\partial \bar{p}}.$$

This is equivalent to

$$\begin{aligned} \left(\frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p} \partial \bar{w}} + \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{w}^2} \frac{\partial \bar{w}}{\partial \bar{p}}\right) \frac{\partial \bar{w}}{\partial \bar{p}} &> \underbrace{\frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p}^2}}_{\leq 0} + 2 \underbrace{\frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p} \partial \bar{w}} \frac{\partial \bar{w}}{\partial \bar{p}}}_{\leq 0} + \underbrace{\frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}} \frac{\partial^2 \bar{w}}{\partial \bar{p}^2}}_{< 0} + \underbrace{\frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{w}^2} \left(\frac{\partial \bar{w}}{\partial \bar{p}}\right)^2}_{\leq 0} \\ \Leftrightarrow \quad 0 &> \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p}^2} + \frac{\partial^2 \pi(\bar{p}, \bar{w})}{\partial \bar{p} \partial \bar{w}} \frac{\partial \bar{w}}{\partial \bar{p}} + \frac{\partial \pi(\bar{p}, \bar{w})}{\partial \bar{w}} \frac{\partial^2 \bar{w}}{\partial \bar{p}^2} \end{aligned}$$

which using the results from above is true and thus establishes the result.

### Proof of Proposition 13: Comparative Statics $\gamma$

We define

$$Y_A \equiv \left(\frac{\partial \pi(\bar{p}_A, \bar{w}_A)}{\partial \bar{p}_A} + \frac{\partial \pi(\bar{p}_A, \bar{w}_A)}{\partial \bar{w}_A} \frac{\partial \bar{w}_A}{\partial \bar{p}_A}\right), \quad Y_B \equiv \left(\frac{\partial \pi(\bar{p}_B, \bar{w}_B)}{\partial \bar{p}_B} + \frac{\partial \pi(\bar{p}_B, \bar{w}_B)}{\partial \bar{w}_B} \frac{\partial \bar{w}_B}{\partial \bar{p}_B}\right),$$

which allows us to rewrite equation (13) as

$$F_{AB} \equiv Y_A \frac{\partial \bar{p}_A}{\partial p_L} + Y_B \frac{\partial \bar{p}_B}{\partial p_L} = 0 \quad \Rightarrow Y_B = -Y_A \frac{\partial \bar{p}_A}{\partial p_L} \left( \frac{\partial \bar{p}_B}{\partial p_L} \right)^{-1}$$

Then the effect of a change in  $\gamma$  on  $p_L$  is given by

$$\frac{dp_L}{d\gamma} = - \frac{\frac{dF_{AB}}{d\gamma}}{\frac{dF_{AB}}{dp_L}}$$

Again, as  $\frac{dF_{AB}}{dp_L} < 0$ ,  $sign\left(\frac{dp_L}{d\gamma}\right) = sign\left(\frac{dF_{AB}}{d\gamma}\right)$

$$\frac{dF_{AB}}{d\gamma} = \underbrace{\frac{\partial Y_A}{\partial \bar{p}_A} \frac{\partial \bar{p}_A}{\partial \gamma} \frac{\partial \bar{p}_A}{\partial p_L} + \frac{\partial Y_B}{\partial \bar{p}_B} \frac{\partial \bar{p}_B}{\partial \gamma} \frac{\partial \bar{p}_B}{\partial p_L}}_{=Part 1} + \underbrace{Y_A \frac{\partial^2 \bar{p}_A}{\partial p_L \partial \gamma} + Y_B \frac{\partial^2 \bar{p}_B}{\partial p_L \partial \gamma}}_{=Part 2}$$

We consider Part 1 and Part 2 separately.

*Part 1* We know that

$$\underbrace{\frac{\partial Y_A}{\partial \bar{p}_A}}_{\leq 0} \underbrace{\frac{\partial \bar{p}_A}{\partial \gamma}}_{< 0} \underbrace{\frac{\partial \bar{p}_A}{\partial p_L}}_{> 0} + \underbrace{\frac{\partial Y_B}{\partial \bar{p}_B}}_{\leq 0} \underbrace{\frac{\partial \bar{p}_B}{\partial \gamma}}_{> 0} \underbrace{\frac{\partial \bar{p}_B}{\partial p_L}}_{> 0}$$

The terms  $\frac{\partial \bar{p}_i}{\partial p_L}$  and  $\frac{\partial \bar{p}_i}{\partial \gamma}$  can be determined through the Implicit Function Theorem, where similar to before we define

$$|J| = \frac{\partial F_A}{\partial p_A} \frac{\partial F_B}{\partial p_B} - \frac{\partial F_A}{\partial p_B} \frac{\partial F_B}{\partial p_A}, \quad |J_i(p_L)| = \frac{\partial F_i}{\partial p_L} \frac{\partial F_{-i}}{\partial p_{-i}} - \frac{\partial F_i}{\partial p_{-i}} \frac{\partial F_{-i}}{\partial p_L}, \quad |J_i(\gamma)| = \frac{\partial F_i}{\partial \gamma} \frac{\partial F_{-i}}{\partial p_{-i}} - \frac{\partial F_i}{\partial p_{-i}} \frac{\partial F_{-i}}{\partial \gamma}$$

Then,  $\frac{\partial \bar{p}_A}{\partial p_L} = -\frac{|J_A(p_L)|}{|J|}$ ,  $\frac{\partial \bar{p}_B}{\partial p_L} = -\frac{|J_B(p_L)|}{|J|}$  and  $\frac{\partial \bar{p}_A}{\partial \gamma} = -\frac{|J_A(\gamma)|}{|J|}$ ,  $\frac{\partial \bar{p}_B}{\partial \gamma} = -\frac{|J_B(\gamma)|}{|J|}$ .

We are interested in comparing the magnitudes of the different effects and in order to compare them we define

$$X_i \equiv 1 + \frac{\partial \hat{g}(\bar{p}_i; \alpha_i)}{\partial \bar{p}_i} (\lambda p_L + (1 - \lambda)(\gamma \bar{p}_i + (1 - \gamma)\bar{p}_j)) - (1 - \lambda)\gamma (1 - \hat{g}(\bar{p}_i; \alpha_i)) > 0$$

We first show that  $\frac{\partial \bar{p}_A}{\partial p_L} < \frac{\partial \bar{p}_B}{\partial p_L}$  which is equivalent to showing that  $-|J_A(p_L)| < -|J_B(p_L)|$ .

$$\begin{aligned} -|J_A(p_L)| &= X_B \lambda (1 - \hat{g}(\bar{p}_A; \alpha_A)) + \lambda(1 - \lambda)(1 - \gamma) (1 - \hat{g}(\bar{p}_A; \alpha_A)) (1 - \hat{g}(\bar{p}_B; \alpha_B)) > 0 \\ -|J_B(p_L)| &= X_A \lambda (1 - \hat{g}(\bar{p}_B; \alpha_B)) + \lambda(1 - \lambda)(1 - \gamma) (1 - \hat{g}(\bar{p}_A; \alpha_A)) (1 - \hat{g}(\bar{p}_B; \alpha_B)) > 0 \end{aligned}$$

Thus  $-|J_A(p_L)| < -|J_B(p_L)|$  if

$$\begin{aligned} &X_B (1 - \hat{g}(\bar{p}_A; \alpha_A)) < X_A (1 - \hat{g}(\bar{p}_B; \alpha_B)) \\ \Leftrightarrow &(1 - \hat{g}(\bar{p}_B; \alpha_B)) \left( 1 + \frac{\partial \hat{g}(\bar{p}_A; \alpha_A)}{\partial \bar{p}_A} (\lambda p_L + (1 - \lambda)(\gamma \bar{p}_A + (1 - \gamma) \bar{p}_B)) \right) \\ &> (1 - \hat{g}(\bar{p}_A; \alpha_A)) \left( 1 + \frac{\partial \hat{g}(\bar{p}_B; \alpha_B)}{\partial \bar{p}_B} (\lambda p_L + (1 - \lambda)(\gamma \bar{p}_B + (1 - \gamma) \bar{p}_A)) \right) \end{aligned}$$

We know that  $(1 - \hat{g}(\bar{p}_B; \alpha_B)) > (1 - \hat{g}(\bar{p}_A; \alpha_A))$  and so it remains to show that

$$\frac{\partial \hat{g}(\bar{p}_A; \alpha_A)}{\partial \bar{p}_A} (\lambda p_L + (1 - \lambda)(\gamma \bar{p}_A + (1 - \gamma) \bar{p}_B)) > \frac{\partial \hat{g}(\bar{p}_B; \alpha_B)}{\partial \bar{p}_B} (\lambda p_L + (1 - \lambda)(\gamma \bar{p}_B + (1 - \gamma) \bar{p}_A))$$

which is guaranteed for  $\gamma \geq \frac{1}{2}$  and a concave  $\hat{g}(\bar{p}_i; \alpha_i)$ . Thus, we have shown that  $\frac{\partial \bar{p}_A}{\partial p_L} < \frac{\partial \bar{p}_B}{\partial p_L}$ .

Next, we show that  $\frac{\partial \bar{p}_B}{\partial \gamma} > -\frac{\partial \bar{p}_A}{\partial \gamma}$ . As before, it suffices to compare  $-|J_A(\gamma)|$  and  $|J_B(\gamma)|$ .

$$\begin{aligned} -|J_A(\gamma)| &= (1 - \hat{g}(\bar{p}_A; \alpha_A))(1 - \lambda)(\bar{p}_B - \bar{p}_A) (X_B - (1 - \lambda)(1 - \gamma)(1 - \hat{g}(\bar{p}_B; \alpha_B))) \\ |J_B(\gamma)| &= (1 - \hat{g}(\bar{p}_B; \alpha_B))(1 - \lambda)(\bar{p}_B - \bar{p}_A) (X_A - (1 - \lambda)(1 - \gamma)(1 - \hat{g}(\bar{p}_A; \alpha_A))). \end{aligned}$$

This reduces to exactly the same condition as before and so we have established that  $\frac{\partial \bar{p}_B}{\partial \gamma} > -\frac{\partial \bar{p}_A}{\partial \gamma}$ . Thus whenever  $\frac{\partial Y_A}{\partial \bar{p}_A}$  is either larger or not much smaller than  $\frac{\partial Y_B}{\partial \bar{p}_B}$ , Part 1 is negative.

A sufficient condition is  $\frac{\partial^3 \pi(\bar{p}, \alpha \hat{f}(\bar{p}; \alpha))}{\partial \bar{p}^3} \leq 0$

*Part 2* We write

$$Y_A \frac{\partial^2 \bar{p}_A}{\partial p_L \partial \gamma} + Y_B \frac{\partial^2 \bar{p}_B}{\partial p_L \partial \gamma} = Y_A \left\{ \frac{\partial^2 \bar{p}_A}{\partial p_L \partial \gamma} - \frac{\partial \bar{p}_A}{\partial p_L} \left( \frac{\partial \bar{p}_B}{\partial p_L} \right)^{-1} \frac{\partial^2 \bar{p}_B}{\partial p_L \partial \gamma} \right\}. \quad (20)$$

We know from the first order condition that  $Y_A$  and  $Y_B$  have opposite signs. As  $Y_i$  is decreasing in  $\bar{p}_i$  and  $\bar{p}_A < \bar{p}_B$ , it has to be the case that  $Y_A$  is positive and  $Y_B$  is negative. Thus, we focus on

$$\text{sign} \left( \frac{\partial^2 \bar{p}_A}{\partial p_L \partial \gamma} - \frac{\partial \bar{p}_A}{\partial p_L} \left( \frac{\partial \bar{p}_B}{\partial p_L} \right)^{-1} \frac{\partial^2 \bar{p}_B}{\partial p_L \partial \gamma} \right) = \text{sign} \left( \frac{\partial \bar{p}_B}{\partial p_L} \frac{\partial^2 \bar{p}_A}{\partial p_L \partial \gamma} - \frac{\partial \bar{p}_A}{\partial p_L} \frac{\partial^2 \bar{p}_B}{\partial p_L \partial \gamma} \right)$$

as  $\frac{\partial \bar{p}_B}{\partial p_L} > 0$ .

We know that  $\frac{\partial \bar{p}_A}{\partial p_L} = \frac{-|J_A(p_L)|}{|J|}$ , which leads to

$$\frac{\partial^2 \bar{p}_A}{\partial p_L \partial \gamma} = \frac{|J| \frac{\partial |J_A(p_L)|}{\partial \gamma} - |J_A(p_L)| \frac{\partial |J|}{\partial \gamma}}{|J|^2}$$

We can then write

$$\begin{aligned} & \frac{\partial \bar{p}_B}{\partial p_L} \frac{\partial^2 \bar{p}_A}{\partial p_L \partial \gamma} - \frac{\partial \bar{p}_A}{\partial p_L} \frac{\partial^2 \bar{p}_B}{\partial p_L \partial \gamma} \\ &= \frac{|J_B(p_L)|}{|J|} \frac{|J| \frac{\partial |J_A(p_L)|}{\partial \gamma} - |J_A(p_L)| \frac{\partial |J|}{\partial \gamma}}{|J|^2} - \frac{|J_A(p_L)|}{|J|} \frac{|J| \frac{\partial |J_B(p_L)|}{\partial \gamma} - |J_B(p_L)| \frac{\partial |J|}{\partial \gamma}}{|J|^2} < 0 \\ &\Leftrightarrow |J_B(p_L)| \frac{\partial |J_A(p_L)|}{\partial \gamma} - |J_A(p_L)| \frac{\partial |J_B(p_L)|}{\partial \gamma} < 0 \end{aligned} \quad (21)$$

We have shown that  $|J_B(p_L)| > |J_A(p_L)| > 0$ . We thus focus on:

$$\begin{aligned} \frac{\partial |J_A(p_L)|}{\partial \gamma} &= \frac{\partial \hat{g}(\bar{p}_B; \alpha_B)}{\partial \bar{p}_B} \lambda(1-\lambda)(1-\hat{g}(\bar{p}_A; \alpha_A))(\bar{p}_B - \bar{p}_A) \\ &\quad - 2\lambda(1-\lambda)(1-\hat{g}(\bar{p}_A; \alpha_A))(1-\hat{g}(\bar{p}_B; \alpha_B)) < 0 \\ \frac{\partial |J_B(p_L)|}{\partial \gamma} &= \frac{\partial \hat{g}(\bar{p}_A; \alpha_A)}{\partial \bar{p}_A} \lambda(1-\lambda)(1-\hat{g}(\bar{p}_B; \alpha_B))(\bar{p}_A - \bar{p}_B) \\ &\quad - 2\lambda(1-\lambda)(1-\hat{g}(\bar{p}_A; \alpha_A))(1-\hat{g}(\bar{p}_B; \alpha_B)). \end{aligned}$$

Equation (21) is true whenever  $\frac{\partial |J_B(p_L)|}{\partial \gamma} \geq 0$ . If  $\frac{\partial |J_B(p_L)|}{\partial \gamma} < 0$ , then it remains true if  $\frac{\partial |J_A(p_L)|}{\partial \gamma} < \frac{\partial |J_B(p_L)|}{\partial \gamma}$  which is equivalent to

$$\frac{\partial \hat{g}(\bar{p}_B; \alpha_B)}{\partial \bar{p}_B} \lambda(1-\lambda)(1-\hat{g}(\bar{p}_A; \alpha_A))(\bar{p}_B - \bar{p}_A) < \frac{\partial \hat{g}(\bar{p}_A; \alpha_A)}{\partial \bar{p}_A} \lambda(1-\lambda)(1-\hat{g}(\bar{p}_B; \alpha_B))(\bar{p}_A - \bar{p}_B),$$

which we already established above.

Based on this we can then calculate the effect of a change in  $\gamma$  on  $\bar{p}_A$  and  $\bar{p}_B$ , taking into account that we know how  $p_L$  adjusts:

$$\begin{aligned} \frac{\partial \bar{p}_A}{\partial \gamma} &= -\frac{1}{|J|} \left( -\lambda \frac{\partial p_L}{\partial \gamma} + (\bar{p}_B - \bar{p}_A)(1-\lambda) \right) (X_B - (1-\lambda)(1-\gamma)(1-\hat{g}(\bar{p}_B; \alpha_B))) < 0 \\ \frac{\partial \bar{p}_B}{\partial \gamma} &= -\frac{1}{|J|} \left( -\lambda \frac{\partial p_L}{\partial \gamma} - (\bar{p}_B - \bar{p}_A)(1-\lambda) \right) (X_A - (1-\lambda)(1-\gamma)(1-\hat{g}(\bar{p}_A; \alpha_A))) \geq 0 \end{aligned}$$

### Proof of Proposition 14: Leader Exclusion

Suppose there are two types of group members,  $A$  and  $B$  and that the share of type  $A$  members is given by  $\kappa$ . Consider without loss of generality the case where the leader excludes type  $A$  members and only group members of type  $B$  belong to the group. The payoff to the group leader is then given by

$$\Pi^B = (1 - \kappa) \pi(\bar{p}_B^B, \alpha_B f(\bar{H}_B^B)),$$

where  $\bar{p}_B^B$  is the type  $B$  steady state identity determined by the optimal leader identity in case the group only contains  $B$  types. We denote the corresponding leader identity by  $p_L^B$ . The steady state is therefore characterized by

$$\bar{p}_B^B = (1 - \hat{g}(\bar{p}_B^B; \alpha_B)) (\lambda_B p_L^B + (1 - \lambda_B) \bar{p}_B^B)$$

Now consider the payoff if agents of type  $A$  are part of the community and the leader chooses his identity  $p_L^{AB}$  such that  $\bar{p}_B^{AB} = \bar{p}_B^B$ . With  $p^{\max}$  sufficiently high, this is always possible by adjusting  $p_L^{AB}$  appropriately. Then the payoff is given by

$$\begin{aligned} \Pi^{AB} &= \kappa \pi(\bar{p}_A^{AB}, \alpha_A f(\bar{H}_A^{AB})) + (1 - \kappa) \pi(\bar{p}_B^{AB}, \alpha_B f(\bar{H}_B^{AB})) \\ &= \kappa \pi(\bar{p}_A^{AB}, \alpha_A f(\bar{H}_A^{AB})) + (1 - \kappa) \pi(\bar{p}_B^B, \alpha_B f(\bar{H}_B^B)) \\ &> (1 - \kappa) \pi(\bar{p}_B^B, \alpha_B f(\bar{H}_B^B)) \end{aligned}$$

Therefore, the payoff when agents of type  $A$  belong to the group is higher than the payoff when they do not belong to the group. Thus, a leader does not have an incentive to omit one type from the group.

## B Skill Investment with Forward-Looking Agents

We discuss optimal skill investment of group members with foresight in two scenarios based on different degrees of foresight: In Scenario 1, group members evaluate their skills investment on a forward-looking basis, considering in particular that skills may be durable. However they remain myopic about future identity changes. In Scenario 2, group members incorporate a prediction of their own future identity as well as the path of other group member identities in the skills investment decision.

### Scenario 1: Group Members Ignore Identity Changes

Consider first a group member with starting identity  $p^0$ , ability  $\alpha$  and discount factor  $\rho$ . The group member takes  $p^0$  as given and fixed forever.

We denote the decay factor of skills by  $\delta \in [0, 1)$  and generalize the cost function to  $C(h, p)$  to bring out the impact of the linearity assumption made in the main part of the paper. The stock of skills  $H$  then acts as a state variable.

The problem of the group member is to maximize:

$$V(H^0) = \sum_{t=1}^{\infty} \rho^t [\alpha f(H^t) - C(H^t - \delta H^{t-1}, p^0)].$$

By the principle of optimality, we can rewrite the problem recursively using the Bellman equation

$$V(H) = \max_{H'} \{ \alpha f(H') - C(H' - \delta H) + \rho V(H') \}.$$

As the period payoff function and the updating function are continuous and bounded we can analyze the problem using standard dynamic programming techniques that yield the Euler equation connecting current investment  $H$ , tomorrow's  $H'$  and for two periods ahead  $H''$ :

$$\alpha f'(H') - C'(H' - \delta H) = -\rho \delta C'(H'' - \delta H') < 0.$$

The right hand side is negative as  $C' > 0$ . Thus, at the optimum a group member invests more today, anticipating that a higher  $H'$  tomorrow will yield additional benefits in the

future by reducing the need to invest then. At the steady state  $H = H' = H'' = \bar{H}$  and thus:

$$\alpha f'(\bar{H}) - C'((1 - \delta)\bar{H}) = -\rho\delta C'((1 - \delta)\bar{H}).$$

The term on the right is zero if  $\rho = 0$  or  $\delta = 0$ . Thus, if skills decay fully and  $\delta = 0$  then myopic outcome is equal to the forward looking one. However, if  $\delta > 0$ , then skills are durable and thus valuable in the future. As a consequence, a group member with  $\rho > 0$  will invest more today than with  $\rho = 0$ . Steady state investment is higher under forward looking behavior than in the myopic case.

## Scenario 2: Group Members Predict Identity Changes

In this section we analyze a variant of our model in which group members are forward-looking and anticipating both their future identity as well as the path of identities of other members of the group. We characterize a forward-looking equilibrium in which each group member solves a dynamic problem maximizing payoffs while taking as given the path of other group members.

**Maximization Problem** Denote by  $\{H_i^t\}_{t=0}^{\infty}$  and  $\{p_i^t\}_{t=0}^{\infty}$  the path of player  $i$ 's skill level and identity, respectively, and similarly  $\{p_{-i}^t\}_{t=0}^{\infty}$  for all other players. Group member  $i$  maximizes:

$$\max_{\{p_i^t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \rho^t \{ \alpha f(H_i^t) - (c_0 + c_1 p_i^t) H_i^t \}$$

subject to the updating equation:

$$p_i^{t+1} = [1 - g(H_i^t; \alpha_i)] \left\{ \lambda p_L + (1 - \lambda) \left[ \gamma_{ii} p_i^t + \sum_{j \neq i} \gamma_{ij} p_j^t \right] \right\}$$

Player  $i$  takes the full path of identities for other players  $\{p_{-i}^t\}_{t=0}^{\infty}$  as given and chooses  $\{p_i^t\}_{t=0}^{\infty}$  accordingly. In equilibrium, all players will individually select a path that represents a best response to other group member's paths. This equilibrium reflects an *open loop* Nash equilibrium and can be interpreted as the Nash equilibrium of the game where players can commit to their future path at time zero.<sup>47</sup>

**Effect of Additional Periods on Skills Investment** To bring out the effect that extending

<sup>47</sup>For further discussion of open loop equilibria and related concepts see e.g. Section 13.4.1 in [Fudenberg and Tirole \(1991\)](#).

the investment horizon has on group member investment decisions when they take account of changes in identity both of themselves and other group members, consider first a finite time problem with final period  $T$ . The myopic decision problem then corresponds to the final period in which group member  $i$  takes  $p_i^T$  and  $p_{-i}^T$  as given and maximizes:

$$\alpha_i f(H_i^T) - (c_0 + c_1 p_i^T) H_i^T$$

As shown in the main part of the paper, the solution is characterized by the first order condition

$$\alpha_i f'(H_i^T) - c_0 + c_1 p_i^T = 0$$

The optimal level of investment is independent of other group members' identities or investment levels.

Now, add one further time period by considering the problem from period  $T-1$ . Denote by  $V_i^T(p_i^T)$  the maximum utility achieved by player  $i$  in period  $T$ . Note that in this notation  $V^T(\cdot)$  explicitly keeps track of time as the problem is not stationary: the value of starting with identity level  $p_i$  in final period  $T$  is different from starting in period  $T-1$ . Furthermore the identities of other players are permitted to change and this means that until a stationary outcome in which identities of other players do not change is reached, we have to keep track of the time period and the corresponding identities for other players in our calculations.

Note further that the value function is decreasing in  $p_i^T$ : a lower identity reduces the marginal cost of skills in investment and thereby allows  $i$  to achieve a higher payoff from investment. The effect can be expressed using the envelope theorem:

$$\begin{aligned} V_i^T(p_i^T) &= \max_{H_i^T} \{ \alpha_i f(H_i^T) - (c_0 + c_1 p_i^T) H_i^T \} \\ (V_i^T)'(p_i^T) &= -c_1 H_i^T < 0 \end{aligned}$$

Consider now the problem from period  $T-1$ . The value function for the group member can then be written as:

$$V_i^{T-1}(p_i^{T-1}) = \max_{H_i^{T-1}} \left\{ \alpha_i f(H_i^{T-1}) - (c_0 + c_1 p_i^{T-1}) H_i^{T-1} + \rho V_i^T(p_i^T) \right\}$$

The first order condition for the choice of  $H_i^{T-1}$  now includes the effect it has on the starting

identity for final period  $T$  via the transition equation.

$$\alpha_i f'(H_i^{T-1}) - (c_0 + c_1 p_i^{T-1}) H_i^{T-1} + \frac{\partial p_i^T}{\partial H_i^{T-1}} \rho(V_i^T)'(p_i^T) = 0 \quad (22)$$

where

$$\frac{\partial p_i^T}{\partial H_i^{T-1}} = -g'(H_i^{T-1}) \left[ \lambda p_L + (1 - \lambda) \sum_j \gamma_{ij} p_j^{T-1} \right] < 0$$

Compared to the myopic approach and the final period, this condition reflects an additional incentive to invest in  $H_i^{T-1}$  represented in the second term on the right hand side of equation (22): a higher  $H_i^{T-1}$  implies player  $i$  will enter the final period with a lower identity and this lower identity means a greater payoff in that final period. Thus, compared to the myopic approach where this effect is ignored, this leads to greater levels of skills and lower identity. The strength of the additional effect depends on the identities of all group members that player  $i$  is connected to. The higher the identities of those other group members, the more they push player  $i$ 's identity up and therefore the greater the payoff from investing in  $H_i^{T-1}$  to shift influence away from the group and towards the host society.

Consider now the derivative of  $V_i^{T-1}(p_i^{T-1})$ . As before, by the envelope theorem we can focus on the direct effect of changes in  $p_i^{T-1}$  yielding the following expression:

$$(V_i^{T-1})'(p_i^{T-1}) = -c_1 H_i^{T-1} + \frac{\partial p_i^T}{\partial p_i^{T-1}} \rho(V_i^T)'(p_i^T) \quad (23)$$

where

$$\frac{\partial p_i^T}{\partial p_i^{T-1}} = [1 - g(H_i^{T-1})] (1 - \lambda) \gamma_{ii} > 0$$

The additional term captures the effect that a change in  $p^{T-1}$  has on next period identity through the identity updating equation.

Turning now to the problem as seen from period  $T - 2$ , the first order condition takes

account of this additional term as follows:

$$\begin{aligned}
& \alpha_i f'(H_i^{T-2}) - (c_0 + c_1 p_i^{T-2}) H_i^{T-2} + \frac{\partial p_i^{T-1}}{\partial H_i^{T-2}} \rho (V_i^{T-1})'(p_i^{T-1}) \\
& = \alpha_i f'(H_i^{T-2}) - (c_0 + c_1 p_i^{T-2}) H_i^{T-2} + \frac{\partial p_i^{T-1}}{\partial H_i^{T-2}} \rho \left\{ -c_1 H_i^{T-1} + \frac{\partial p_i^T}{\partial p_i^{T-1}} \rho (V_i^T)'(p_i^T) \right\} \\
& = 0
\end{aligned}$$

As both terms in curly braces are negative, there is thus an additional incentive for player  $i$  to choose a higher level of  $H_i^{T-2}$  relative to time period  $T - 1$ : a higher skill level implies a lower next period identity that – in addition to reducing marginal cost of investment in period  $T - 1$  – will also lead to a lower identity in period  $T$ , adding further benefits. Note that the strength of this effect depends on the parameter  $\gamma_{ii}$ , i.e., the extent to which past identity influences next period identity for player  $i$ .

As we roll back the problem further in time, these benefits will accumulate further, so that in period 1 an increase in  $H_i^1$  will reduce not only  $p_i^2$  but also all future  $p_i^t$ . Note that overall the problem remains well defined due to concavity of  $f(\cdot)$  and the linearity of marginal costs. There thus exists a unique solution of the sequence problem for any given path of other group members' identities over time.

**Forward-Looking Skills Investment at the Steady State** As noted above, for a given path of identities of other group members the optimization problem of group member  $i$  is not necessarily stationary across time. To be able to say more about impact of forward-looking behavior on long-run skills investment we consider a group which has reached a steady state.

In steady state  $p_i^t = \bar{p}_i$ ,  $H_i^t = \bar{H}_i$  as well as  $V^t(p_i^t) = V(\bar{p}_i)$  for all players and thus we have a stationary investment problem from the perspective of group member  $i$ . This allows dropping the superscript  $t$  from the value functions and the derivative of the value function in equation (23) can then be rearranged to yield

$$\begin{aligned}
V'(\bar{p}_i) &= \frac{-c_1 \bar{H}_i}{1 - \rho(1 - \lambda) [1 - g(\bar{H}_i)] \gamma_{ii}} \\
&= (-c_1 \bar{H}_i) \sum_{t=0}^{\infty} [\rho(1 - \lambda) [1 - g(\bar{H}_i)] \gamma_{ii}]^t
\end{aligned}$$

The expression in terms of an infinite sum highlights the way in which an increase in identity  $p_i$  creates benefits in terms of marginal cost reductions proportional to  $c_1 H_i$  both directly

and indirectly via its effect on next period  $p_i$  and subsequent periods for the infinite future.

The first order condition for the optimal skills level in the steady state can be rearranged to

$$\begin{aligned} & \alpha_i f'(\bar{H}_i) - (c_0 + c_1 \bar{p}_i) \\ & = -\rho \frac{c_1 \bar{H}_i}{1 - \rho(1 - \lambda) [1 - g(\bar{H}_i)] \gamma_{ii}} g'(\bar{H}_i) \left[ \lambda p_L + (1 - \lambda) \left( \gamma_{ii} \bar{p}_i + \sum_{j \neq i} \gamma_{ij} \bar{p}_j \right) \right] \end{aligned}$$

This condition shows that relative to the myopic optimization problem, the group members have an incentive to invest more in skills, resulting in lower steady state identities. This incentive derives from the negative relationship between skills and identity and the extent to which lower identity reduces the marginal costs of further investment, both directly, and indirectly through lowering future identities via the self-influence parameter  $\gamma_{ii}$ .

### Implications for Convergence of Identities

The derivations in the previous sections bring out the effect of extending the planning horizon of group members in their decision how much to invest in skills. As we show, making players forward-looking tends to lead to greater investment levels than the myopic case, both in the setting with durable skills (Scenario 1) and in the setting with players predicting future paths for their own and other players' identities (Scenario 2). Through the function  $g(\cdot)$  these higher investment levels will translate mechanically into an overall lower path of group member identities compared to the myopic case.

In terms of implications for the dynamics of our model, we can capture these differences by suitably adjusting the function  $\hat{g}(p)$ , which maps a group member's identity into the weight assigned to the host society and implicitly reflects the optimal choice of investment. With forward looking agents we would expect to see this function to increase for every  $p$ . However, as long as the resulting new function  $\hat{g}(p)$  satisfies our Assumption 1, the convergence of the system overall is assured as before and Proposition 3 will go through. Furthermore, the qualitative properties of the long run outcome such as incomplete assimilation of group members remain unaffected.

# School of Economics and Finance



**This working paper has been produced by  
the School of Economics and Finance at  
Queen Mary University of London**

**Copyright © 2016 Anja Prummer and Jan-Peter Siedlarek.  
All rights reserved**

**School of Economics and Finance  
Queen Mary University of London  
Mile End Road  
London E1 4NS**

**Tel: +44 (0)20 7882 7356**

**Fax: +44 (0)20 8983 3580**

**Web: [www.econ.qmul.ac.uk/research/workingpapers/](http://www.econ.qmul.ac.uk/research/workingpapers/)**