working paper 15 17

Community Leaders and the Preservation of Cultural Traits

Anja Prummer and Jan-Peter Siedlarek

FEDERAL RESERVE BANK OF CLEVELAND

Working Paper 15-17 October 2015

Community Leaders and the Preservation of Cultural Traits

Anja Prummer and Jan-Peter Siedlarek

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 Integration, Cultural Transmission, Leadership, Immigrants, Labor Market Outcomes, Social Influence, Networks

JEL Classification: J15, Z10, D02.

Suggested citation: Prummer, Anja, and Jan-Peter Siedlarek, 2015. "Community Leaders and the Preservation of Cultural Traits," Federal Reserve Bank of Cleveland, working paper no. 15-17.

Anja Prummer is at Cambridge-INET Institute (ap809@cam.ac.uk) and Jan-Peter Siedlarek is at the Federal Reserve Bank of Cleveland (jan-peter.siedlarek@ clev. frb.org. The authors are grateful to 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. They also thank seminar participants at Oxford University, University of Cambridge, SITE Stockholm School of Economics, NYU, INET-New York, CTN 2014, ASREC 2014, SAET 2014 and EEA 2014. Siedlarek acknowledges financial support through DFG grant SFB-TR15.

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.[1] 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 the assimilation process 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 can 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

[1] As an example, gains from assimilation are higher earnings if an immigrant conforms to the norms prevalent in the host country, which then 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.[2]

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)).[3] 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-

[2]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)).

[3]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.[4] 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. But different 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 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.

[4]The identity dynamics provide a departure from Akerlof and Kranton (2000, 2010), where identity remains fixed.

Then, assimilation will remain incomplete. In the presence of a leader, each period 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, individuals that 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 lead 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.[5]

Second, we look at Muslim assimilation, which has been argued to be different from that of other ethnic and religious groups.[6] 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 parents' preferences for cultural traits (Bisin and Verdier (2000)),[7] 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). One key

[5]For an overview of the Jewish assimilation, see Carvalho and Koyama (2011).

[6]For an overview, see Bisin et al. (2008), Constant et al. (2006), and Haug (2008).

[7]Fernández et al. (2004) and Fernández and Fogli (2006) also show that cultural traits are shaped by parents.

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. Full assimilation occurs even if parents have preferences for the persistence of the cultural trait. 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 transition 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 transmission 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.[8] Büchel et al. (2014) prove that their model converges with a time-varying transmission 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 is not applicable to our settings. We discuss these technical aspects in more detail below Proposition 3.

Our technical contribution here 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 but instead exploits a different condition restricting the speed of change of the transition matrix as well as the specific social configuration with 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 transmission 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.[9] 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

[8] 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 transition matrix.

[9] Other work on opinion dynamics includes Acemoglu et al. (2010), Acemoglu et al. (2013), Büchel et al. (2012), and Lorenz (2006).

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 (2013) discuss the role of cultural leaders in the clash of cultures. In their settings 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.[10]

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 in Section 5. Section 6 connects our findings to a number of applications. Section 7 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.[11] The host society does not take any actions 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}.[12] We fix $p_S^t = p_S = 0$ for the host society for all time periods.[13] We also take the leader's identity

[10] Empirically, Munshi and Wilson (2008) have also emphasized the importance of churches in the preservation of cultural traits.

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

[12] 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.

[13] 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.

as fixed for all t once it has been chosen, $p_L^t = p_L \in [0, p^{max}]$.[14] Thus, $p_i^t = 0$ indicates full identification by group member i with the host country and $p_i^t = p_L$ 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.

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. 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.[15] 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.

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

(iii) Group Members The weight group member i assigns to group member j is denoted by $\gamma_{ij}(1-\lambda)(1-g(H_i^{t-1}))$, 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

[14] It will turn out that in the long run having a fixed identity p_L will be optimal for the leader even if we allow them to change it every period in principle.

[15] 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.

of their social connections.

An overview of the process is provided in Figure 1. The identity adaptation process for

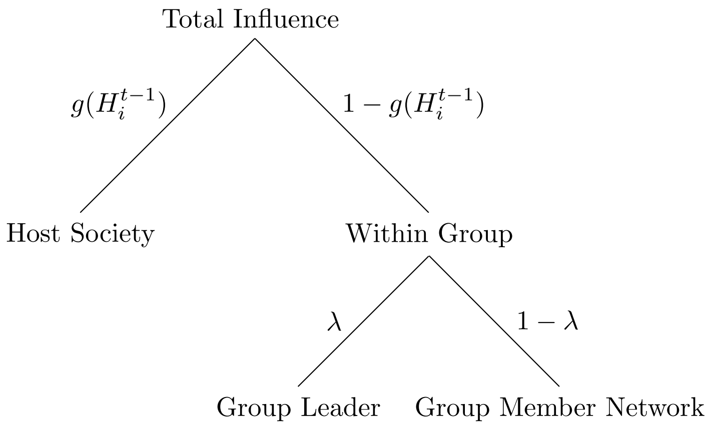

Figure 1: Influence Weights in Period t

group member i from p_i^{t-1} to p_i^t can then be summarized as follows,

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

where we have used the fact that $p_S = 0$ so that the influence term of the host society vanishes. Note that the normalization of the host society's identity to zero is without loss of generality.

To illustrate the workings of within group influence, we consider the symmetric case of two group members who each assign the same weight $\gamma \in [0,1]$ to themselves and $(1-\gamma)$ on the other group member. Equation (1) then simplifies to

$$p_i^t = \left[1 - g\left(H_i^{t-1}\right)\right] \cdot \left\{\lambda p_L + (1-\lambda)\left(\gamma p_i^{t-1} + (1-\gamma) p_j^{t-1}\right)\right\}.$$

Setting $\gamma = 1$ reflects the special case of the group member being isolated from the influence of other group members.

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 labour market.[16]

Each period the group members simultaneously select how much to invest in skills.

[16]Skills should not be interpreted as schooling or human capital that is useful both in the immigrant and the native labour market, but rather as skills that only give an advantage in the native labor market compared to the immigrant one.

The payoff from skills H_i^t is given by $\alpha_i f\left(H_i^t\right)$. 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 $c\left(p_i^t\right)$, where $c'\left(p_i^t\right) > 0$ so that these are strictly increasing in p_i^t, 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. Additionally, note that past and current investment reduces future investment due to the features of the identity adaptation process.

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 a payoff function that depends on the identity and the wealth or earnings, $\alpha f(H)$, of the group members as follows:

$$\sum_{i=1}^{n} \pi(p_i^t, \alpha_i f(H_i^t)) \tag{2}$$

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. We further assume that

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

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

Our assumptions on the payoff function reflect the discussion of community leader motivations in the introductory section: leaders care about the identity of their community as well as the its financial wellbeing. Religious leaders may care about the latter from a need to maximize donations (Ceylan (2010); Bekkers and Wiepking (2010)) or indirectly from a sense of prestige or altruism. Foreign-language media can better sell their products to a wealthy community that still culturally identifies with its origin culture and is thus interested in what it has to offer.

Motivated by these examples, 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 (3) and (4):

$$\pi_{RL}(p_i^t, H_i^t) = \beta p_i^t + (1-\beta)\alpha_i f(H_i^t), \tag{3}$$

$$\pi_{BL}(p_i^t, H_i^t) = \left(p_i^t\right)^{\beta} \left(\alpha_i f(H_i^t)\right)^{1-\beta}, \tag{4}$$

Equations (3) and (4) 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 $\sigma \to 0$. The parameter σ governs the extent to which the leader sees identity and earnings as strategic substitutes. Therefore, the business leader, who only cares about identity as it enhances his earnings, considers the two as strategic complements, for the religious leader they are neither. In the latter, we follow Carvalho and Koyama (2011) who use exactly the same payoff function for Rabbis. We consider these payoff functions to be examples that serve to highlight the intuition behind our general results, but none of our main theoretical results depends on these specific functional forms.

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.

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.[17] The optimal investment is given by the solution of the following maximization problem:

$$\max_{H_i^t \geq 0} \quad \left\{ \alpha_i f\left(H_i^t\right) - c\left(p_i^t\right) H_i^t \right\}. \tag{5}$$

This set-up assumes that group members are myopic. We allow in the Appendix for two types of forward looking agents and show that this assumption does not affect our qualitative results. We solve the maximization problem by a first order approach, which 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 α_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 α_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 α_i as a higher α_i implies a higher $H^*(p_i^t; \alpha_i)$, which leads to a 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

[17] 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.

identity levels as follows:

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

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

$$\overline{p}_i = \left[1 - \hat{g}\left(\overline{p}_i; \alpha_i\right)\right] \left\{ \sum_{j=1}^{n} \gamma_{ij} \overline{p}_j \right\}. \tag{7}$$

The corresponding levels of steady state investment $\overline{H}_i$ is equivalent to the optimal skill investment, $H^*(\overline{p}_i; \alpha_i)$. 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.[18]

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 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 $\hat{g}(p_i^t; \alpha_i)$.[19]

[18] The result is unsurprising and mirrors D'Amico et al. (2009)'s analysis of uni-reducible Markov chains.

[19] Additionally, even if some group members might not find it beneficial to invest in skills for a sufficiently high level of identity, there will be full assimilation without a leader as long as they are connected and thus influenced by another group member, who finds it beneficial to invest in skills. Then their identity will still adjust until they also invest in skills. Thus, in the long run, there will be full assimilation, as long as each group member who is not investing initially is connected to someone who does invest.

We then turn to the convergence process. We can discuss the speed at which group members assimilate in isolation of other group members (i.e. assuming $\gamma_{i,i} = 1 \ \forall \ i$) as a function of their initial identity and their productivity. 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. The result is driven by the assumption that $g(H_i)$ is strictly increasing in H and summarized in the following proposition.

Proposition 2 (Speed of Convergence). *Assume $\gamma_{i,i} = 1 \ \forall \ i$ and take two group members l and k.*

1. *If both group members have the same ability ($\alpha_k = \alpha_l$) and the initial identity of group member k is lower than that of l ($p_k^0 < p_l^0$), then the identity of k converges faster than that of l.*
2. *If both k and l have the same initial identity ($p_k^0 = p_l^0$) and k has higher productivity than l ($\alpha_k > \alpha_l$), then the identity of k converges faster than that of l.*

3.2 Identity Dynamics with Leader

We now introduce a leader with some fixed p_L and it will turn out that the group leader indeed will choose a fixed identity in the long run. We do this to obtain the steady state identities of the group members. Based on these steady state identities, the leader then sets the optimal identity.[20]

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

$$p_i^{t+1} = \left[1 - \hat{g}\left(p_i^t; \alpha_i\right)\right] \left[\lambda p_L + (1-\lambda) \sum_{j=1}^{n} \gamma_{ij} p_j^t\right]. \tag{8}$$

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.

$$\left|\frac{\partial \hat{g}(p_i; \alpha_i)}{\partial p_i}\right| \frac{p_i}{\hat{g}(p_i; \alpha_i)} < \frac{p_i}{p^{max}} \qquad \forall \ p_i, \alpha_i \tag{9}$$

[20] 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.

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 the their group does not decline too rapidly in response to small decreases in their identity p_i.

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

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

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. [21] Specifically, Büchel et al. (2014) require the transition 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.[22] Panebianco (2014) studies a model with multiple traits and vertical as well as oblique socialization. His convergence result for the case with time-varying socialization matrix (Corollary 1) requires strictly positive diagonal entries and a socialization pattern with a single communication class in each component. These conditions are not satisfied in our case as the model we study consists of a single component and two communication classes: both the leader and the host society form a separate communication class. Furthermore, we allow for zero diagonal entries

[21] 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.

[22] 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.

amongst group members, which violates Assumption 1 in Panebianco (2014).[23]

In addition to dealing with a case not studied elsewhere, we offer a novel approach to establish convergence and uniqueness of 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 multiple steady states.

In terms of our application, the result shows that the presence of a leader, who is not susceptible to any influence, guarantees that there is no longer full assimilation.[24] However, the group members will assimilate to some extent in *any* steady state. If the steady state is unique, we can study the comparative statics of the extent of assimilation. In particular, the steady state identity $\overline{p}_i$ is given by

$$\overline{p}_i = [1 - \hat{g}(\overline{p}_i; \alpha_i)] \left\{ \lambda p_L + (1-\lambda) \sum_{j=1}^{n} \gamma_{ij} \overline{p}_j \right\} \tag{10}$$

and $\overline{H}_i = H^*(\overline{p}_i; \alpha_i)$.

Effect of α and γ We study the effect of α as well as the effect on γ of steady state identities as foundation for our analysis of policy instruments. If all group members have the same α_i, that is the same productivity, then all group members will respond symmetrically and thus have the same $\hat{g}(\overline{p}_i; \alpha_i)$ function. As a consequence the long run identities of the group members are the same, independently of the level of connectedness γ. This insight readily carries over to any social structures.

Remark 4 (Ability $\alpha_i = \alpha \; \forall \; i$)**.**
If all group members have the same level of ability, the steady state identity vector is invariant to the structure of the group network.

[23] To see that our setting is not covered in Panebianco (2014), consider the following example with only one group member. Let the updating matrix alternate between T_{odd} and T_{even}, where $T_{\text{odd}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ .3 & .2 & .5 \end{pmatrix}$ and $T_{\text{even}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ .2 & .3 & .5 \end{pmatrix}$. Then, it is straightforward to verify that the process does not converge.

[24] Again, we will show, that it is optimal for the leader to not be influenced.

In contrast, if there is heterogeneity in ability α_i, then the long run identities of the group members differ. To see this more clearly, suppose 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 γ specifying the total weight put on all group members of the same type.[25] Steady state identity levels for A and B in this setting can then be summarized by:

$$\overline{p}_A = [1 - \hat{g}(\overline{p}_A; \alpha_A)] \{\lambda p_L + (1-\lambda) [\gamma \overline{p}_A + (1-\gamma) \overline{p}_B]\} \tag{11}$$

$$\overline{p}_B = [1 - \hat{g}(\overline{p}_B; \alpha_B)] \{\lambda p_L + (1-\lambda) [\gamma \overline{p}_B + (1-\gamma) \overline{p}_A]\} \tag{12}$$

Note that for a given identity level A invests more in skills than B resulting in a higher weight on the host society for A. Based on equations (11) and (12) we can then analyze steady state behavior of this setting in Proposition 5.

Proposition 5 (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 $\overline{p}_A < \overline{p}_B$.*
2. *A lower group cohesion, that is a higher γ 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 $\frac{\partial \overline{p}_A}{\partial \gamma} < 0$ and $\frac{\partial \overline{p}_B}{\partial \gamma} > 0$. This implies that if group cohesion is low, 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 will consider the strategic choice a leader faces and we show that the leaders benefit from fixing a position. Leaders have an incentive to choose norms and values that result in incomplete assimilation, but do not benefit from adjusting their identity in every

[25]Note that as group members of the same type are homogeneous, the within-type influence weights do not affect the dynamics and final outcome.

period.[26] The reason of why assimilation fails is thus that the leader does not find it optimal to be influenced by the host society or group members.

Leader Identity Choice with Homogenous Group Members For the leader identity choice we focus on payoffs in the steady state and we further assume that $\alpha_i = \alpha$ 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 $\overline{p}$ and group members' earnings $\alpha f(\overline{H})$. The steady state skill investment is again determined by the steady state identity and so we rewrite income as $\alpha \hat{f}(\overline{p}; \alpha) \equiv \alpha f(H^*(\overline{p}; \alpha))$. Note that $\hat{f}(\overline{p}; \alpha)$ is *decreasing* in $\overline{p}$. This derives from the fact that $f(\overline{H})$ is increasing in $\overline{H}$ and $H^*(\overline{p}; \alpha)$ is decreasing in $\overline{p}$.

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

$$\overline{p} = [1 - \hat{g}(\overline{p}; \alpha)] \{\lambda p_L + (1 - \lambda) \overline{p}\}, \tag{13}$$

Given Assumption 1, the impact of p_L on $\overline{p}$ is strictly positive.[27] This implies that a higher identity of the leader is always associated with lower assimilation of group members. Thus, when a leader chooses greater cultural distinction, the immigrant community will be less assimilated. We then determine under what circumstances the leader chooses full assimilation (by setting $p_L = 0$), becomes an extremist (setting $p_L = p^{max}$) or chooses intermediate assimilation ($0 < p_L < p^{max}$).

We define:

$$MRS \equiv \frac{\pi_1(\overline{p}, \alpha \hat{f}(\overline{p}, \alpha))}{\pi_2(\overline{p}, \alpha \hat{f}(\overline{p}; \alpha))} \tag{14}$$

as the marginal rate of substitution between identity and earnings for the leader. Similarly,

$$MRT \equiv -\alpha \frac{\partial \hat{f}(\overline{p}; \alpha)}{\partial \overline{p}} \tag{15}$$

gives the marginal rate of transformation at which the leader can transform identity into earnings, where an increase in identity leads to lower earnings.

[26] In particular, they do not want to adjust identity in the long run, in steady state, but prefer to keep their position fixed.

[27] See the Appendix for a formal proof.

Proposition 6 (Leader Decision)**.** *The leader decision depends on the relationship of MRS and MRT.*

- Full Assimilation: *If $MRS < MRT$ for any $p \in [0, p^{max}]$, then the leader sets $p_L = 0$ implementing full assimilation.*
- Extremism: *If $MRS > MRT$ for any $p \in [0, p^{max}]$, then the leader sets $p_L = p^{max}$, the highest possible cultural distinction.*
- Intermediate Assimilation: *If there exists a $\overline{p}$ such that $MRS = MRT$ and $\hat{f}(\overline{p}; \alpha)$ is a concave function, then the leader implements an intermediate level of assimilation and $p_L \in (0, p^{max})$.*

In summary, which outcome emerges depends on the specific functional forms involved. There might not be an interior solution, that is, there does not exists a $p \in [0, p^{max}]$ such that the MRS is equal to the MRT. In this case, the leader chooses either full assimilation or extremism.

If an interior solution exists and the group member's payoff function is a concave function in $\overline{p}$, then intermediate cultural distinction emerges.[28] 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 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, depending on the properties of $\hat{f}(p; \alpha)$. To shed some more light on the intuition behind these results, we turn to the special cases of the religious and business leader.

Corollary 6.1 (Religious Leader)**.** *If an interior solution exists and $\hat{f}(p; \alpha)$ is a convex function, then the religious leader selects either the most extreme norms and values, p^{max}, or full assimilation.*

If the earnings function is convex, then 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

[28]The concavity of the earnings function is driven by the sign of $f'''(H)$. We provide further details in the Appendix.

extreme identity yields the highest payoff. This depends on p^{max} as well as on α and β. If p^{max} is sufficiently large, then the religious leader always prefers the highest possible cultural distinction. For smaller p^{max}, a higher productivity α and lower β, where β is the weight the leader assigns to identity, make full assimilation more likely.

Corollary 6.2 (Business Leader)**.** *The business leader will never choose full assimilation.*

Similar to the religious leader, the business leader might also choose the most extreme level of norms. 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.

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 α, and last the effect of a lump sum transfer, z, the effect of a change in group ties γ.

Effect of Change in Productivity We are interested in the effect of an increase in α, 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 α sheds light on the impact of a policy aimed at decreasing discrimination.[29] Our finding is summarized in Proposition 7.

Proposition 7 (Comparative Statics α)**.** .
If the leader chooses intermediate assimilation, then as α increases:

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

In case of intermediate assimilation, a higher α results in an increase in the identity the leader sets. An increase in α 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

[29] 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.

assimilated group members. Note that identity and skill investment are not only affected 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 7.1 (Religious vs Business Leader: Skill Investment)**.** .
With a religious leader skill investment is unaffected by a change in α, 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

$$\beta + (1-\beta)\alpha f'(\overline{H})\frac{\partial \overline{H}}{\partial \overline{p}} = 0 \qquad \Leftrightarrow \qquad \beta + (1-\beta)\frac{f'(\overline{H})}{f''(\overline{H})}c_1 = 0 \tag{16}$$

Equation (16) is independent of α and so the optimal H^* is unaffected by a change in α. In contrast, for the business leader a change in α 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 α without the presence of a leader induces immigrants to invest more in skill, 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 effect of an increase in α 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 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.

We again restrict attention to homogenous group members and set $\alpha_i = \alpha$ for all i. The income of a group member is then given by $\alpha \hat{f}(\overline{p};\alpha) + z$

Proposition 8 (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 9 (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 $\overline{p}(p^{max})$ 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

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

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 α, 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 α or a lump sum transfer z on welfare. It is natural to assume that welfare depends both on earnings and identity.[30] Then an increase in α makes individuals better off (assuming costs are

[30] Alternatively, welfare could only depend on earnings and then the effect of a parameter change on welfare is the same as that on earnings.

not too high), as *both* earnings and identity increase following an increase in either α. Similarly, an increase in z makes individuals better off. Thus, these type of policies increase the welfare of immigrants, but they do not foster cultural integration.[31]

Heterogenous Group Members So far, we have restricted attention to the case of homogenous group members with the same α. But we are also interested in the case, when α differs and in particular, in the way the optimal identity set by the leader responds to group members being more or less connected. We consider again a simplified setting with two different types, A and B where $\alpha_A > \alpha_B$, and symmetric influence structure summarized by Equations (11) and (12), and study the effect of a change in γ. We interpret γ here as a measure of "group cohesion", that is, how important ties are between members within the group. Having a lower γ implies that an individual is 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 of how assimilationist policies might be structured for different types of groups.

Proposition 10 (Comparative Statics γ)**.** *Consider the two-type setting with $\alpha_A > \alpha_B$ characterized in Equations (11) and (12). Let $\hat{g}(\overline{p_i}; \alpha_i)$ be a concave function and $\gamma \geq \frac{1}{2}$. Additionally, assume that $\frac{\partial \pi(\overline{p}, \alpha \hat{f}(\overline{p}; \alpha))}{\partial \overline{p}}$ is concave. Suppose the leader chooses intermediate assimilation. Then, a decrease in γ, that is an increase in group cohesion leads to a greater level of cultural distinction, to a higher p_L. The identity of A decreases and so his investment and earnings increase. The effect on identity and thus investment and earnings of B is ambiguous.*

One implication of $\hat{g}(\overline{p_i}; \alpha_i)$ being concave is that a decrease in the weight that group members assign to each other will lead to a decrease in $\overline{p}_A$ that is smaller than the increase $\overline{p}_B$. So, if γ increases then the leader has an incentive to reduce p_L. To see this more clearly, consider the first order condition of the leader when faced with a group that consists of two types:

$$\left(\frac{\partial \pi(\overline{p}_A, \overline{w}_A)}{\partial \overline{p}_A} + \frac{\partial \pi(\overline{p}_A, \overline{w}_A)}{\partial \overline{w}_A} \frac{\partial \overline{w}_A}{\partial \overline{p}_A} \right) \frac{\partial \overline{p}_A}{\partial p_L} + \left(\frac{\partial \pi(\overline{p}_B, \overline{w}_B)}{\partial \overline{p}_B} + \frac{\partial \pi(\overline{p}_B, \overline{w}_B)}{\partial \overline{w}_B} \frac{\partial \overline{w}_B}{\partial \overline{p}_B} \right) \frac{\partial \overline{p}_B}{\partial p_L} = 0 \tag{17}$$

[31] It is straightforward to incorporate an explicit welfare analysis as the impact of the α and z on earnings and identity go in the same direction.

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

$$\left(\frac{\partial \pi(\overline{p}_A, \overline{w}_A)}{\partial \overline{p}_A} + \frac{\partial \pi(\overline{p}_A, \overline{w}_A)}{\partial \overline{w}_A}\frac{\partial \overline{w}_A}{\partial \overline{p}_A}\right) > 0 > \left(\frac{\partial \pi(\overline{p}_B, \overline{w}_B)}{\partial \overline{p}_B} + \frac{\partial \pi(\overline{p}_B, \overline{w}_B)}{\partial \overline{w}_B}\frac{\partial \overline{w}_B}{\partial \overline{p}_B}\right) \tag{18}$$

This implies that when γ increases, for fixed p_L, equation (17) will be negative. The leader can now increase or decrease p_L, which depends also on how much p_L affects $\overline{p}_A$ and $\overline{p}_B$. But we also know that a change in p_L affects $\overline{p}_B$ more than $\overline{p}_A$ as B assigns more weight to the leader compared to A. This implies that the leader will find it optimal to decrease p_L in response. Put differently, if γ decreases he finds it best to increase p_L, implying that societies with a higher group cohesion will induce the leader to choose a higher cultural distinction.

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

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

of group members A and B is not too large.[32]

The decrease in p_L following an increase in γ leads for group member A to a decrease in identity. His identity decreases when γ increases and p_L remains fixed. Now, p_L decreases as well leading to a further decrease 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 increase if γ decreases and p_L were kept fixed. The decrease in p_L now goes against this effect and so the overall affect on identity is ambiguous. The same holds true for the effect on his skill investment and earnings.

6 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

[32] The condition is always fulfilled for the religious leader if $f^{(4)}(H) \leq 0$.

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.[33] 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 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, whilst 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 were 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 were strong anti-Semitism emerged in Germany, see Katz (1980).[34] 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

[33] 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.

[34] 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.

imply significant differences in α 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.[35]

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

[35] For evidence on this and a summary of the situation of the Jewish communities, see Carvalho and Koyama (2011).

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.[36] In terms of our model this translates into a lower β, 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 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

[36] See an article by The Guardian on www.theguardian.com/commentisfree/2007/jun/05/vacancyforanimam

correlation between extremism, terrorism and poor economic conditions.[37]

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 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 labour 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.[38]

[37] 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 α. 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 α and are omitted here.

[38] Targeting the income of 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.

7 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 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. An exploratory analysis of data from the UK Citizenship Survey suggests a negative relationship between notions of feeling "British" and geographic distance to places of worship for Muslim communities. This implies that if a Muslim lives closer to a Mosque, his values differ more from those of other UK citizens compared to a Muslim living further away. This holds even for individuals who are religious, but have not based their residence choice on religion. The driving force behind this result is also not the fact that living close to a Mosque results in more Muslim friends, which might be an indication that the religious institution per se matters. However, we are doubtful that these results indeed provide sufficient evidence for our theory. The fact that the presence and characteristics of leaders may be endogenous to the group itself severely limits the identification of causal relationships between leader and integration outcome and we have therefore chosen to not include our empirical work.

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.

Acemoglu, D. and J. Robinson (2012). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Publishing Group.

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.

Aleksynska, M. and Y. Algan (2010). Assimilation and Integration of Immigrants in Europe. Technical report, Dsicussion Paper Forschungsinstitut zur Zukunft der Arbeit.

Algan, Y., A. Bisin, A. Manning, and T. Verdier (2012). *Cultural Integration of Immigrants in Europe*. Oxford University Press.

Algan, Y. and P. Cahuc (2006). Job Protection: The Macho Hypothesis. *Oxford Review of Economic Policy 22*(3), 390–410.

Algan, Y., C. Dustmann, A. Glitz, and A. Manning (2010). The Economic Situation of First and Second-Generation Immigrants in France, Germany and the United Kingdom. *The Economic Journal 120*(542), F4–F30.

Austen-Smith, D. and R. G. Fryer (2005). An Economic Analysis of "Acting White". *The Quarterly Journal of Economics 120*(2), 551–583.

Bala, V. and S. Goyal (1998). Learning from Neighbours. *The Review of Economic Studies 65*(3), 595–621.

Battu, H. and Y. Zenou (2010). Oppositional Identities and Employment for Ethnic Minorities: Evidence from England. *The Economic Journal 120*(542), F52–F71.

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.

Bekkers, R. and P. Wiepking (2011). Who Gives? A Literature Review of Predictors of Charitable Giving Part One: Religion, Education, Age and Socialisation. *Voluntary Sector Review 2*(3), 337–365.

Berger, I. E. (2006). The Influence of Religion on Philanthropy in Canada. *Voluntas: International Journal of Voluntary and Nonprofit Organizations 17*(2), 110–127.

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., E. Patacchini, T. Verdier, and Y. Zenou (2010). Bend It Like Beckham: Ethnic Identity and Integration. National Bureau of Economic Research Working Paper No w16465.

Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou (2011a). Ethnic Identity and Labour Market Outcomes of Immigrants in Europe. *Economic Policy 26*(65), 57–92.

Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou (2011b). Formation and Persistence of Oppositional Identities. *European Economic Review 55*(8), 1046–1071.

Bisin, A. and G. Topa (2003). Empirical Models of Cultural Transmission. *Journal of the European Economic Association 1*(2-3), 363–375.

Bisin, A., G. Topa, and T. Verdier (2004). Religious Intermarriage and Socialization in the United States. *Journal of Political Economy 112*(3), 615–664.

Bisin, A., G. Topa, and T. Verdier (2009). Cultural Transmission, Socialization and the Population Dynamics of Multiple-Trait Distributions. *International Journal of Economic Theory 5*(1), 139–154.

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. (1992). Ethnic Capital and Intergenerational Mobility. *The Quarterly Journal of Economics 107*(1), 123–150.

Borjas, G. (2013). The Slowdown in the Economic Assimilation of Immigrants: Aging and Cohort Effects Revisited Again. Working Paper 19116, National Bureau of Economic Research.

Borjas, G. and R. Freeman (1992). *Introduction and Summary, Economic Consequences for the U.S. and Source Areas*. University of Chicago.

Borjas, G. J. (1985). Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants. *Journal of Labor Economics 3*(4), pp. 463–489.

Borjas, G. J. (1995). Ethnicity, Neighborhoods, and Human-Capital Externalities. *The American Economic Review 85*(3), pp. 365–390.

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, S. Klößner, et al. (2012). Opinion Dynamics under Conformity. Working Paper 469, Bielefeld University, Institute of Mathematical Economics.

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.

Carvalho, J.-P. and M. Koyama (2013). Resisting Education. unpublished.

Casey, T. and C. Dustmann (2010). Immigrants' Identity, Economic Outcomes and the Transmission of Identity Across Generations. *The Economic Journal 120*(542), F31–F51.

Cavalli-Sforza, L. L. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press.

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.

Chiswick, B. and P. Miller (1992). *Language in the Immigrant Labor Market in Immigration, Language and Ethnicity: Canada and the United States*. American Enterprise Institute, Washington.

Chiswick, B. R. (1991). Speaking, Reading, and Earnings among Low-Skilled Immigrants. *Journal of Labor Economics 9*(2), 149–170.

Constant, A., L. Gataullina, K. Zimmermann, and L. Zimmermann (2006). Clash of Cultures: Muslims and Christians in the Ethnosizing Process. Available at SSRN.

Constant, A. F. and K. F. Zimmermann (2008). Measuring Ethnic Identity and its Impact on Economic Behavior. *Journal of the European Economic Association 6*(2-3), 424–433.

D'Amico, G., J. Janssen, and R. Manca (2009). The Dynamic Behaviour of Non-Homogeneous Single-unireducible Markov and Semi-Markov Chains. In *Networks, Topology and Dynamics*, pp. 195–211. Springer.

Danzer, A. M. and F. Yaman (2010). Ethnic Concentration and Language Fluency of Immigrants in Germany. SOEPpaper.

DeGroot, M. (1974). Reaching a Consensus. *Journal of the American Statistical Association 69*(345), 118–121.

DeMarzo, P., D. Vayanos, and J. Zwiebel (2003). Persuasion Bias, Social Influence, and Unidimensional Opinions. *The Quarterly Journal of Economics 118*(3), 909–968.

Eguia, J. X. (2013). Discrimination and Assimilation. Available at SSRN 1657144.

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.

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. (1998). *We Are All Multiculturalists Now*. Harvard University Press.

Glazer, N. and D. P. Moynihan (1963). *Beyond the Melting Pot: The Negroes, Puerto Ricans, Jews, Italians and Irish of New York City.* Massachusetts Inst. Technology Pres.

Golub, B. and M. Jackson (2010). Naive Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics 2*(1), 112–149.

Grenier, G. (1984). The Effects of Language Characteristics on the Wages of Hispanic-American Males. *Journal of Human Resources 19*(1), 35–52.

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. Stichs (2012). Islamisches Gemeindeleben in Deutschland. Technical report, Bundesamt für Migration und Flüchtlinge and Stiftung Zentrum für Türkeistudien und Integrationsforschung.

Haug, S. (2008). *Sprachliche Integration von Migranten in Deutschland*. Bundesamt für Migration und Flüchtlinge.

Hauk, E. and H. Mueller (2013, September). Cultural leaders and the clash of civilizations. *Journal of Conflict Resolution*.

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.

Horst, U., A. Kirman, and M. Teschl (2006). Changing Identity: The Emergence of Social Groups. Working Paper 2006-51, GREQAM Groupement de Recherche en Economie Quantitative d'Aix-Marseille.

Judd, K. L. (1998). *Numerical Methods in Economics*. MIT press.

Kantarevic, J. (2004). Interethnic Marriages and Economic Assimilation of Immigrants. Discussion paper series, 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.

Kossoudji, S. A. (1988). English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men. *Journal of Labor Economics 6*(2), 205–228.

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.

McManus, W. (1990). Labor Market Effects of Language Enclaves: Hispanic Men in the United States. *Journal of Human Resources 25*(2), 228–252.

McManus, W., W. Gould, and F. Welch (1983). Earnings of Hispanic Men: The Role of English Language Proficiency. *Journal of Labor Economics 1*(2), 101–130.

McManus, W. S. (1985). Labor Market Costs of Language Disparity: An Interpretation of Hispanic Earnings Differences. *The American Economic Review 75*(4), 818–827.

Meng, X. and R. G. Gregory (2005). Intermarriage and the Economic Assimilation of Immigrants. *Journal of Labor economics 23*(1), 135–174.

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.

Nielsen, H. S., M. Rosholm, N. Smith, and L. Husted (2004). Qualifications, Discrimination, or Assimilation? An Extended Framework for Analysing Immigrant Wage Gaps. *Empirical Economics 29*(4), 855–883.

Panebianco, F. (2014). Socialization Networks and the Transmission of Interethnic Attitudes. *Journal of Economic Theory 150*, 583–610.

Pichler, M. (2010a). Cultural Formation of Preferences and Assimilation of Cultural Groups. Working Paper 438, Bielefeld University, Institute of Mathematical Economics.

Pichler, M. (2010b). The Economics of Cultural Formation of Preferences. Working Paper 431, Bielefeld University, Institute of Mathematical Economics.

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.

Reich, B. (2012). Culture, Community and Segregation. unpublished.

Rizzo, H., A.-H. Abdel-Latif, and K. Meyer (2007). The Relationship between Gender Equality and Democracy: A Comparison of Arab versus non-Arab Muslim Societies. *Sociology 41*(6), 1151–1170.

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.

Schervish, P. G. (1993). The Dependent Variable of the Independent Sector: the Definition and Measurement of Giving and Volunteering. *Voluntas: International Journal of Voluntary and Nonprofit Organizations 4*(2), 223–232.

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.

Thornton, A., D. F. Alwin, and D. Camburn (1983). Causes and Consequences of Sex-Role Attitudes and Attitude Change. *American Sociological Review 48*(2), 211–227.

Vaughan, D. (2013). Nurture vs. Nurture: Endogenous Parental and Peer Effects and the Transmission of Culture. Working Paper 2013-04, Working Papers, Banco de México.

Vella, F. (1994). Gender Roles and Human Capital Investment: The Relationship Between Traditional Attitudes and Female Labour Market Performance. *Economica 61*(242), 191–211.

Wiepking, P. and R. Bekkers (2012). Who Gives? A Literature Review of Predictors of Charitable Giving. Part Two: Gender, Family Composition and Income. *Voluntary Sector Review 3*(2), 217–245.

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

Zaleski, P. A. and C. E. Zech (1992). Determinants of Contributions to Religious Organizations. *American Journal of Economics and Sociology 51*(4), 459–472.

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.

APPENDIX

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

Verify by inspection that $\overline{p}_i = 0 \; \forall \; i$ is a solution to the steady state condition in Equation (7). To establish uniqueness and convergence, rewrite Equation (6) in matrix form as follows:

$$p^{t+1} = \left[I - \hat{G}\left(p^t, \alpha\right)\right] \Gamma p^t, \tag{19}$$

where p^t is the vector of identities. I is the identity matrix and we define $\hat{G}\left(p^t, \alpha\right)$ to be the $n \times n$ matrix with diagonal elements $\hat{G}_{ii}\left(p^t\right) = \hat{g}(p_i^t; \alpha_i)$ and zeros elsewhere. Γ is the group member influence matrix defined by $\Gamma_{ij} = \gamma_{ij}$. Solving backwards yields:[39]

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

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

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

as required.

Proof of Proposition 2 – Speed of Convergence

Recall the one-step updating Equation 6 with γ_{ii} set to one:

$$p_i^{t+1} = \left[1 - \hat{g}(p_i^t; \alpha_i)\right] p_i^t. \tag{20}$$

The first part of the statement follows from the fact that both $\left[1 - \hat{g}(p_i^t; \alpha_i)\right]$ and p_i^t are strictly increasing in p_i^t. Thus for any pair of starting of identities $p_k^0 < p_l^0$ we have $p_k^1 < p_l^1$. Then, by a simple induction argument it follows that $p_k^t < p_l^t \; \forall \; t \geq 0$, so that k is always strictly more assimilated than l as required.

[39] For this proof a superscript t applied to a matrix such as Γ denotes Γ taken to the tth power whilst a superscript on a vector such as p refers to the vector p in period t.

The second part of the result follows analogously, using the fact that $\hat{g}(p_i^t; \alpha_i)$ is strictly increasing in α_i.

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 (8). 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

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

which is strictly greater than zero because $\hat{g}(\tilde{p}_i; \alpha_i) < 1$, $\lambda > 0$ 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

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

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 contraction.

Uniqueness and Convergence

We show that the updating process with leader described by Equation (8) 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 (8) 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:[40]

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

We use is the matrix norm induced by the ∞-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

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

we derive the following sufficient condition:

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

[40]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 λ (in particular $\lambda \to 0$). This condition further simplifies to:

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

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 5: Identity Spread for Different Types

The results follow from the implicit function theorem applied to the system of equations describing the steady state (Equations (11) and (12)). We define

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

By the implicit function theorem the effect of a change in the parameter γ 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 θ. We compute $|J|$ and $|J_A(\gamma)|$ from the partial derivatives given by:

$$\frac{\partial F_i}{\partial \overline{p}_i} = [1 - \hat{g}(\overline{p}_i; \alpha_i)](1-\lambda)\gamma - \frac{\partial \hat{g}(\overline{p}_i; \alpha_i)}{\partial p_i} \{\lambda p_L + (1-\lambda)\left[\gamma \overline{p}_i + (1-\gamma)\overline{p}_j\right]\} - 1 < 0 \quad (21)$$

$$\frac{\partial F_i}{\partial \overline{p}_j} = [1 - \hat{g}(\overline{p}_i; \alpha_i)](1-\lambda)(1-\gamma) > 0 \quad (22)$$

$$\frac{\partial F_i}{\partial \gamma} = [1 - \hat{g}(\overline{p}_i; \alpha_i)](1-\lambda)\left(\overline{p}_i - \overline{p}_j\right). \quad (23)$$

Equation (21) can be signed using Assumption 1. The sign of Equation (22) follows directly from $\hat{g}(\overline{p}_i; \alpha_i) \in (0, 1)$. Equation (23) adopts the sign of $(\overline{p}_i - \overline{p}_j)$ 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 γ between types.*

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

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

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

2. *$|\overline{p}_A - \overline{p}_B|$ decreases in γ.*

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

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

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

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

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

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

Proof of Proposition 6: Leader Decision

We first establish that $\overline{p}(p_L)$ is an increasing function in p_L. To see this recall Equation (13)

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

Differentiating with respect to p_L yields

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

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

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

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

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

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

We can have three different cases:

(1) The leader sets $p_L = p^{max}$ if

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

(2) He sets $p_L = 0$ if

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

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

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

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

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

The second order condition is negative whenever:

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

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

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

Thus, a sufficient condition for an interior solution that maximizes leader's payoff is $\frac{\partial^2 \hat{f}(\overline{p};\alpha)}{\partial \overline{p}^2} < 0$ which is implied by $f'''\left(\overline{H}\right) < 0$. 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 $p_L = 0$.

Proof of Proposition 7: Comparative Statics α

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

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

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

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

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

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

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

Further,

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

where

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

Then,

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

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

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

$$\begin{aligned} c_1 \frac{d\overline{p}}{d\alpha} - f'(\overline{H}) = {} & -f'(\overline{H}) \frac{\partial^2 \pi(\overline{p}, \overline{w})}{\partial \overline{p}^2} + \frac{\partial^2 \pi(\overline{p}, \overline{w})}{\partial \overline{p} \partial \overline{w}} \left(-c_1 f(\overline{H}) - f'(\overline{H}) \frac{\partial \overline{w}}{\partial \overline{p}}\right) \\ & + \frac{\partial^2 \pi(\overline{p}, \overline{w})}{\partial \overline{w}^2} \frac{\partial \overline{w}}{\partial \overline{p}} \left(-c_1 f(\overline{H})\right). \end{aligned} \tag{24}$$

The sign of Equation (24) is ambiguous without further restrictions.

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

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

which is equivalent to

$$\frac{f(\overline{H})f''(\overline{H})}{f'(\overline{H})}\frac{\partial F}{\partial \overline{p}} > c_1 \frac{\partial F}{\partial \alpha} + f' \frac{\partial F}{\partial \overline{p}}. \tag{25}$$

We know from Equation (24) that the term on the right can be rewritten as

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

Equation (25) can then be rewritten as

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

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

Proof of Corollary 7.1: Increase in Skill Investment

We show that an increase in α leads to an increase in skill investment for the business leader. By Equation (24):

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

The derivatives are given by:

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

We can then write the statement of the corollary as

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

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

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

This establishes the result.

Proof of Proposition 10: Comparative Statics γ

We define

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

which allows us to rewrite Equation (17) as

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

Then the effect of a change in γ on p_L is given by

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

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

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

We consider Part 1 and Part 2 separately.

Part 1 We know that

$$\underbrace{\frac{\partial Y_A}{\partial \overline{p}_A}}_{\leq 0}\underbrace{\frac{\partial \overline{p}_A}{\partial \gamma}}_{<0}\underbrace{\frac{\partial \overline{p}_A}{\partial p_L}}_{>0}+\underbrace{\frac{\partial Y_B}{\partial \overline{p}_B}}_{\leq 0}\underbrace{\frac{\partial \overline{p}_B}{\partial \gamma}}_{>0}\underbrace{\frac{\partial \overline{p}_B}{\partial p_L}}_{>0} \tag{26}$$

The terms $\frac{\partial \overline{p}_i}{\partial p_L}$ and $\frac{\partial \overline{p}_i}{\partial \gamma}$ 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, $\frac{\partial \overline{p}_A}{\partial p_L}=-\frac{|J_A(p_L)|}{|J|}$, $\frac{\partial \overline{p}_B}{\partial p_L}=-\frac{|J_B(p_L)|}{|J|}$ and $\frac{\partial \overline{p}_A}{\partial \gamma}=-\frac{|J_A(\gamma)|}{|J|}$, $\frac{\partial \overline{p}_B}{\partial \gamma}=-\frac{|J_B(\gamma)|}{|J|}$.

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

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

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

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

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

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

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

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

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

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

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

This reduces to exactly the same condition as before and so we have established that $\frac{\partial \overline{p}_B}{\partial \gamma} > -\frac{\partial \overline{p}_A}{\partial \gamma}$. Thus whenever $\frac{\partial Y_A}{\partial \overline{p}_A}$ is either larger or not much smaller than $\frac{\partial Y_B}{\partial \overline{p}_B}$, Part 1 is negative. A sufficient condition is $\frac{\partial^3 \pi(\overline{p}, \alpha \hat{f}(\overline{p};\alpha))}{\partial \overline{p}^3} \leq 0$

Part 2 We write

$$
Y_A \frac{\partial^2 \overline{p}_A}{\partial p_L \partial \gamma} + Y_B \frac{\partial^2 \overline{p}_B}{\partial p_L \partial \gamma} = Y_A \left\{ \frac{\partial^2 \overline{p}_A}{\partial p_L \partial \gamma} - \frac{\partial \overline{p}_A}{\partial p_L}\left(\frac{\partial \overline{p}_B}{\partial p_L}\right)^{-1} \frac{\partial^2 \overline{p}_B}{\partial p_L \partial \gamma} \right\}. \tag{27}
$$

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

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

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

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

$$
\frac{\partial^2 \overline{p}_A}{\partial p_L \partial \gamma} = \frac{|J|\frac{\partial |J_A(p_L)|}{\partial \gamma} - |J_A(p_L)|\frac{\partial |J|}{\partial \gamma}}{|J|^2} \tag{28}
$$

We can then write

$$
\begin{aligned}
&\frac{\partial \overline{p}_B}{\partial p_L}\frac{\partial^2 \overline{p}_A}{\partial p_L \partial \gamma} - \frac{\partial \overline{p}_A}{\partial p_L}\frac{\partial^2 \overline{p}_B}{\partial p_L \partial \gamma} \\
=&\frac{|J_B(p_L)|}{|J|}\frac{|J|\frac{\partial |J_A(p_L)|}{\partial \gamma} - |J_A(p_L)|\frac{\partial |J|}{\partial \gamma}}{|J|^2} - \frac{|J_A(p_L)|}{|J|}\frac{|J|\frac{\partial |J_B(p_L)|}{\partial \gamma} - |J_B(p_L)|\frac{\partial |J|}{\partial \gamma}}{|J|^2} < 0 \\
\Leftrightarrow &|J_B(p_L)|\frac{\partial |J_A(p_L)|}{\partial \gamma} - |J_A(p_L)|\frac{\partial |J_B(p_L)|}{\partial \gamma} < 0
\end{aligned} \tag{29}
$$

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

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

Equation (29) 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

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

which we already established above.

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

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

Proof of Proposition 8: Effect of a Lump Sum Transfer

This proof follows along the lines of the Proof of Proposition 7, with the only change being the lump sum transfer added to income from skills and thus $\overline{w} = \alpha f(H^*) + z$. The sign of a change in the transfer z is given by $sign\left(\frac{\partial \overline{p}}{\partial z}\right) = sign\left(\frac{\partial F}{\partial z}\right)$, where

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

We know that

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

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

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

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

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

Wages are positive if

$$1 > \frac{\frac{\partial F}{\partial z}}{\frac{\partial F}{\partial \overline{p}}} \frac{\partial \overline{w}}{\partial \overline{p}} \qquad \Leftrightarrow \qquad \frac{\partial F}{\partial z} \frac{\partial \overline{w}}{\partial \overline{p}} > \frac{\partial F}{\partial \overline{p}}.$$

This is equivalent to

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

$$\Leftrightarrow \qquad 0 > \frac{\partial^2 \pi(\overline{p}, \overline{w})}{\partial \overline{p}^2} + \frac{\partial^2 \pi(\overline{p}, \overline{w})}{\partial \overline{p} \partial \overline{w}} \frac{\partial \overline{w}}{\partial \overline{p}} + \frac{\partial \pi(\overline{p}, \overline{w})}{\partial \overline{w}} \frac{\partial^2 \overline{w}}{\partial \overline{p}^2}$$

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

Technical Note: 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 in the skills investment decision.

Scenario 1: Group Members Ignore Identity Changes

Consider first a group member with starting identity p^0, ability α and discount factor ρ. 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:

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

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

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

As the period payoff function and the transition 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'':

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

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 $H = H' = H'' = \overline{H}$ and thus:

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

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.

Finally, note that aside from the higher investment level for any given p, the qualitative features of the analysis mirror that of the myopic case. Higher p leads to lower investment, albeit at different levels. Thus, with $\hat{g}(p;\alpha)$ and Assumption 1 suitably redefined, the steady state results concerning persistence of cultural traits go through as for myopic group members.

Scenario 2: Group Members Predict Identity Changes

In a second model of forward looking group members, the group member takes account of future changes to her own identity as a function of today's actions. However, the prediction leaves out any future changes in other group members' identities. The prediction is thus perfectly rational for an isolated group member, without connections to others. For an agent with additional connections it is only imperfectly rational, as it ignores the changing influence of other group members, when instead they also are updated every period. The group members in effect act as rational, but atomistic, agents.

Consider as above a group member with starting identity p^0, ability α and discount factor ρ. Assume in addition linear costs of skills investment as in the main part of the paper. The problem of the group member is to maximize:

$$V(p_0) = \sum_{t=0}^{\infty} \rho^t \left[\alpha f(H^t) - \left(c_0 + c_1 p^t\right) H^t\right]$$

subject to the transition equation:

$$p^{t+1} = \left[1 - g(H^t;\alpha)\right] \left\{\lambda p_L + (1-\lambda)\left\{\gamma p^t + (1-\gamma)\tilde{p}^0_{others}\right\}\right\},$$

where γ is the influence weight the agent puts on herself and $\tilde{p}^0_{others}$ is the influence weighted average identity of other group members at time zero.

As in Scenario 1 the dynamic programming problem is continuous and bounded. By the principle of optimality we can thus rewrite the problem recursively using the Bellman equation:

$$V(p) = \max_{H} \left\{ \alpha f(H) - (c_0 + c_1 p) H + \rho V(p') \right\}$$

We restate the problem using p' as control. H is then a function of p' by solving the transition equation for H:

$$H(p') = g^{-1} \left(1 - \frac{p'}{\lambda p_L + (1-\lambda) \{\gamma p + (1-\gamma)\tilde{p}_{others}\}} \right)$$

This expression describes the relationship between H and p' answering the question what is the level of H that is required to achieve p' in the next period.[41] $H(p')$ is decreasing in p'. The problem then becomes:

$$V(p) = \max_{p'} \left\{ \alpha f(H(p')) - (c_0 + c_1 p) H(p') + \rho V(p') \right\}.$$

Standard techniques for dynamic programming yield the Euler equation connecting H and H':

$$\alpha f'(H(p')) - (c_0 + c_1 p) = \rho c_1 \frac{H(p'')}{H'(p')} < 0.$$

The right hand side is negative as $H(p'') > 0$ and $H'(p') < 0$. Thus for any given p, at the optimum a group member generates a greater decrease in p' and invests more today, which will yield additional benefits in the future. At the steady state $p = p' = p'' = \overline{p}$ and thus:

$$\alpha f'(H(\overline{p})) - (c_0 + c_1 \overline{p}) = \rho c_1 \frac{H(\overline{p})}{H'(\overline{p})}.$$

Again, the right hand side is negative for a forward-looking group member with $\rho > 0$ whilst for a myopic agent with $\rho = 0$ it is zero. The effect of forward looking behavior on the steady state level of identity $\overline{p}$ is ambiguous. A decrease in $\overline{p}$ reduces the first term on the left ($f'(\cdot)$) but in addition also the marginal cost term by factor c_1. The overall effect can

[41]Note that the domain of $H(p')$ is restricted by the properties of the function $\hat{g}(p;\alpha)$ to the open interval $(0, \lambda p_L + (1-\lambda)\{\gamma p + (1-\gamma)\tilde{p}_{others}\})$.

then be positive or negative.

However, as argued for Scenario 1, with suitable adjustments to Assumption 1, the qualitative results of our analysis concerning the persistent of cultural traits remain unaffected with forward-looking group members.