Working Papers

Institute of
Mathematical
Economics

438

August 2010

Cultural Formation of Preferences and Assimilation of Cultural Groups

Michael M. Pichler

IMW · Bielefeld University
Postfach 100131
33501 Bielefeld · Germany

email: imw@wiwi.uni-bielefeld.de
http://www.wiwi.uni-bielefeld.de/~imw/Papers/showpaper.php?438
ISSN: 0931-6558

Cultural Formation of Preferences and Assimilation of Cultural Groups

Michael M. Pichler*
IMW, Bielefeld University, Universitätsstr. 25, D–33615 Bielefeld
Tel.: +49 521 106 4902, Fax: +49 521 106 2997
mpichler@uni–bielefeld.de
and
eri–ces, University of Valencia

First Draft
August 2010

Abstract

Based on the cultural formation of continuous preferences framework of Pichler [16], this paper analyzes the evolution of preferences and behavior in a two cultural groups setting. We show that the qualitative dynamic properties depend crucially on what parents perceive as the optimal preferences for their children to adopt. Under inter–generationally fixed optimal preferences, the preferences of the cultural groups will always stay distinct. If the optimal preferences coincide with those derived from the representative group behavior, then a multitude of convergence path types can realize. These contain both an inter–generational assimilation process toward the same preference point, as well as inter–generational dissimilation.

Keywords: Continuous Preferences; Assimilation; Cultural Groups; Endogenous Preferences; Preference Evolution; Socialization
JEL–Classification numbers: C72, J13, Z13

*This research project has been generously supported by the Deutsche Forschungsgesellschaft (DFG). The author thanks Walter Trockel, Herbert Dawid and Gonzalo Olcina for invaluable advice. I am further grateful for useful comments from Fernando Louge, Stefan Zeugner, as well as from participants at various workshops and conferences.

1 Introduction

Motivation When different cultural groups live together, then there is always cultural exchange through the social(ization) interactions between the members of the groups. While this can well concern the mutual dissemination of the customs of the groups, it notably consists to a large extent of a mutual (inter–generational) influencing of the preferences, values, norms, attitudes and beliefs of the groups' members.

This context raises interest both on empirical and theoretical grounds. In the empirical context, the question of assimilation and integration of immigrants with different cultural backgrounds into hosting societies has attained increasing attention in recent years, both in media and on the political agenda. This calls for a framework that allows for a theoretical representation and analysis, optimally leading into a leveraged understanding of the empirical processes at work.

The present paper presents such a theoretical framework, based on a recent theory of Pichler [16] on the inter–generational formation of continuous preferences.[1] We will show a static and dynamic analysis of the evolution of behavior and preferences in a two cultural groups setting, subject to one type of continuous preferences. Thereby, one of the focus points will be to derive conclusions about the underlying assimilation process between the two cultural groups, both in terms of their adopted preferences, as well as in terms of their behavioral decisions.

Contributions and Results The first part of this paper is devoted to a recapitulation of the cultural formation of preferences framework of Pichler [16]. In doing so, we will show in a first step how children come to adopt intensities of any arbitrary continuous preference type. We let this be based on the children's social learning from role models for preference intensities observed in their social environment. Thereby, we derive these role models, which we call *displayed preference intensities*, from the socio–economic actions of adults. We then show how to interpret the preference intensities that adults have adopted such as to construct and characterize preferences over displayed preference intensities, respectively the underlying socio–economic actions.

In a second step, we introduce one possible way to endogenize the cultural formation of preference process as resulting out of purposeful parental socialization decisions. These are twofold. The first is the choice of a displayed preference intensity. The second consists of investments into the weight that this role model has in the socialization process of the child relative to the weight that the observed representative displayed preference

[1] The latter are meant to contain all types of preferences that can reflect different intensities, located in a convex subset of the real line.

intensity of the general social environment has. We will call this weight the *parental socialization success share*. Thus, basically, the parental decision problem is to choose best replies against the representative role model of the general social environment. Notably, this is subject to the perception that the parents have of the optimal preference intensity for their children to adopt (and different perceptions can have a remarkable impact on the qualitative static and dynamic properties, as will be shown below). We show conditions under which a pure strategy Nash equilibrium of the induced 'strategic socialization interaction game' of the parents exists. These equilibrium choices do endogneously determine the inter–generational *evolution* of the preference intensities and the preferences of the society.

In the second and main part of this paper, we then embed the endogenous cultural formation of preferences process in a society that is populated by two distinct cultural groups. With these, we basically refer to a collection of families, for which it holds that the parental (adult) members have identical adopted preference intensities and form identical perceptions of the optimal preference intensities for their children. Under these (and other) symmetry assumptions it then follows that all parents of the same cultural group do always have identical best reply choices. This is the basis for the existence of group–symmetric Nash equilibria. Notably, under such choices, (almost) all children of the same cultural group do adopt the same preference intensities, which inter–generationally preserves the symmetry properties.

The central task pursued in this paper is the analysis of the group–symmetric Nash equilibrium choices and the resulting dynamic evolution of the adopted preference intensities under two different benchmark perception rules for the optimal preference intensities. In the main part of the paper, we consider first exogenously fixed (and distinct) optimal preference intensities, and second the case where the parents of a group perceive the average displayed preference intensity of their own group members as the reference value ('endogenous norms'). Finally, in the Appendix, we also discuss the case where all parents have 'imperfect empathy'.[2]

Under any possible perception rule for optimal preference intensities, the direction of the socialization efforts of the parents of both groups is always toward the optimum. In the first case that we consider, this leads to an inter–generational coordination toward a situation where the positions of the adopted preference intensities can be considered 'consistent' with the relative location of the fixed optimal preference intensities (if this situation has not been given initially). With this we mean that (a) the group with the strictly larger fixed optimal preference intensity does also have a strictly larger adopted preference intensity, and (b) the preference intensities of both

[2]In a more general context, Pichler [16] has already shown that this case features a global 'melting pot' property, i.e. the adopted preference intensities of (almost) all dynasties converge to the same point.

groups do lie strictly between the two optima.

Within this 'generic state space', the socialization efforts of the members of the two cultural groups are in the opposite directions. This yields the result that the parents of both cultural groups dis–integrate behaviorally (i.e. the parents with the strictly larger/lower adopted preference intensity choose to display a strictly larger/lower than adopted preference intensity) and choose strictly positive parental socialization success shares. This has the consequence that the relative positions of the two cultural groups are inter–generationally preserved and the 'generic state space' can not be left.

While as a steady state must exist in the 'generic state space' (only), we could even show that under weak conditions, one such rest point must exist which is even asymptotically stable.

The qualitative asymptotic results of the fixed optimal preference intensity case do thus feature the opposite extreme to the 'imperfect empathy' case: While as in the latter case, the preferences of the cultural groups do always converge to the same point, this will never happen under fixed optimal preference intensities.

Compared to this sort of uniqueness of the qualitative asymptotic properties, the case of endogenous norms features a larger variety of possible convergence path types. First of all, we could show that under suitable conditions, any sequence of adopted preference intensities of the two groups converges to a steady state. Even, there is a basin in terms of a maximum distance of the adopted preference intensities, such that all pairs of adopted preference intensities that enter (or start in) this basin converge to a point where all adults have the same preferences. However, for a large enough initial preference distance, it is possible that the cultural groups dissimilate on the transitory path and a steady state with a larger than initial preference distance is reached.

Notably, these latter dynamic results are subject to a normalization of the selection of the group–symmetric Nash equilibrium choices in any given period. Specifically, we consider only choices that inter–generationally preserve the relative positions of the two cultural groups in terms of a lower/larger adopted preference intensity. If the preferences of the two groups are distinct, this is guaranteed in the case where the parents of both cultural groups dis–integrate behaviorally, coupled with a strictly positive choice of socialization success share. In case of identical preferences, the analogous condition is that all parents do not actively socialize, i.e. they choose a displayed preference intensity that coincides with the adopted one, coupled with zero investments into the socialization success share. Thus, any such point is a rest point.

Related Literature The present analysis stands in a close relation to few existing contributions on the question of the cultural formation of continuous preferences. Important early treatments of the topic are Cavalli-Sforza and Feldman [9] in a theoretical, and Otto et al. [14] in an empirical context. More recently Bisin and Topa [1] proposed a representation of the formation of the values of continuous cultural traits. They represented the adopted value of the cultural trait (or continuous preference type) as a weighted average between a role model that is taken by the family and the average of the value of the cultural trait in the population.

The major limitation of this contribution is, however, that it does neither explicitly consider the family's choice of role models, nor the construction of role models themselves. Rather, Bisin and Topa [1] assume that parents always choose their 'target value' (i.e. the optimal preference intensity in the terminology of the present paper) as a role model. Given this restricted view on the family's behavioral choices, its socialization decision is then reduced to the choice of its weight in the formation of the preference intensity of their child.[3]

A second, and well established, related strand is the literature on the economics of cultural transmission. It has been introduced by Bisin and Verdier [3, 4, 5] and Bisin et al. [2], and is based on the work of Cavalli-Sforza and Feldman [8, 9] and Boyd and Richerson [7] in evolutionary anthropology. The focus is on the analysis of the population dynamics of the distribution of a discrete set of preferences (respectively cultural traits) under an endogenous intergenerational cultural transmission mechanism.

The endogeneity stems from the purposeful parental choice of socialization intensity, which effectively determines the probability that the child will directly adopt the preferences of the parents. Parents engage into the cost of purposeful socialization in order to avoid (decrease the probability) that their child will not adopt their preferences — in which case parents encounter subjective utility losses. For an exhaustive overview of foundations of and contributions to this literature, see Bisin and Verdier [6].[4]

Notably, the latter theory considers the probabilistic *transmission* of preferences and does not approach the issue of formation of the latter. This restricts its applicability mainly to *discrete preferences* (respectively cultural traits).

Outline The further setup of this paper is as follows. The succeeding section 2 recapitulates the general framework on the (endogenous) cultural formation of continuous preferences of Pichler [16]. This is followed by the analysis of static and dynamic properties of the model in a two cultural groups setting in section 3. We consider both fixed optimal preference in-

[3] The same sort of critique applies to the approach of Panebianco [15], who considers the formation of inter–ethnic attitudes.

[4] Related to this strand of literature are the contributions of Cox and Stark [10] and Stark [19] on the 'demonstration (or preference shaping) effect' of parental altruism choices in front of their children.

tensities in subsection 3.1, as well as endogenous norms in subsection 3.2. Further, Appendix B.1 contains a short treatment of the case where all parents have 'imperfect empathy'. Finally, Appendix B.2 discusses essential comparative statics properties of the model, and section 4 concludes.

2 Cultural Formation of Preferences

This section discusses a general model of the formation of continuous preferences through the socialization process (in subsection 2.1). In subsection 2.2, we will also show how this cultural formation of preferences process can be derived out of optimal parental socialization decisions. Notably, the framework developed here constitutes a shortened representation of the one introduced in Pichler [16]. For the details, please confer the original source directly. The reader who is familiar with the latter can read the present section as a refresher, but can well directly proceed to section 3.

2.1 Cultural Formation of Preferences

Consider an overlapping generations society. In the present and next subsection, we will restrict our glance on the cultural formation of preferences process between two succeeding generations. This makes it possible to drop all time indexes (for ease of exposition).

In any given period, let our society be populated by a continuum of adults, $a \in A = [0, 1]$ endowed with Lebesgue measure λ, and their children. For simplicity, we will assume that reproduction is asexual and every adult has one offspring, so that we can denote with $\tilde{a} \in \tilde{A}$ the children of the parents $a \in A$ (and the population size is constant).

Let us assume that all adults have available the same feasible set of socio–economic actions, $X \subseteq \mathbb{R}^n$. The structure of the latter is such that any typical element $x \in X$ is the characteristic role model for exactly one preference intensity (PI). We will call this the *displayed preference intensity* (DPI) of a choice of socio–economic actions x, $\phi^d(x) \in \mathbb{R}$. Thus, there exists a displayed preference intensity function

$$\phi^d : X \mapsto \mathbb{R}$$

where $\phi^d(X)$ then corresponds to the set of possible DPIs. To simplify the subsequent exposition, we will denote the DPI of the socio–economic actions of adult $a \in A$, $x_a \in X$, as $\phi^d_a := \phi^d(x_a)$.

We will now introduce the representation of the socialization process that this paper proposes. This will be established on grounds of the *tabula rasa* assumption, which means in the present context that children are born with undefined PI, and equally, with undefined preferences. On this basis, we then let the formation of the PI that a child adopts result out of social learning from the DPIs of adults (only) that it is confronted with.

Specifically, this is being embedded in a framework of socialization inside the family and by the general adult social environment, or 'direct vertical and oblique socialization'.[5] In this context, we will let the PI that a child $\tilde{a} \in \tilde{A}$ adopts be formed according to a weighted average between the representative DPIs of both socialization sources. In the case of the child's family, this coincides with the DPI of its single parent $a \in A$, $\phi_a^d \in \phi^d(X)$. The representative DPI of the child's general social environment, $A_a := A \backslash \{a\}$, will be denoted $\phi_{A_a}^d$. This results out of the child's social learning from the observed DPIs of (eventually) different subsets of adults that it is confronted with.

More precisely, we assume that there is a finite partition of the adult set, $\{A_J\}_{J=1}^K$, and that all children socially learn from the *average* DPIs of these subsets, $\phi_{A_J}^d := \frac{1}{\lambda(A_J)} \int_{A_J} \phi_{a'}^d \, d\lambda\,(a') \in \text{con } \phi^d(X)$, $\forall J = 1, \ldots, K$. Specifically, for every child $\tilde{a} \in \tilde{A}$ there are *oblique socialization weights*, $\{\sigma_{\tilde{a}J}\}_{J=1}^K$, that represent the relative cognitive impacts of the child's social learning from the various subsets of adults. These weights satisfy $\sigma_{\tilde{a}J} \in [0,1]$ and $\sum_{J=1}^K \sigma_{\tilde{a}J} = 1$, $\forall \tilde{a} \in \tilde{A}$, $\forall J = 1, \ldots, K$. We obtain, $\forall \tilde{a} \in \tilde{A}$,

$$\phi_{A_a}^d := \sum_{J=1}^K \sigma_{\tilde{a}J} \phi_{A_J}^d \in \text{con } \phi^d(X).$$

The weight that the DPI of the parent of a child $\tilde{a} \in \tilde{A}$ has in the socialization process of the child will be called the *parental socialization success share*, $\hat{\sigma}_a \in [0,1]$. This corresponds to the cognitive impact of the parental DPI relative to the cognitive impact of the representative DPI of the child's general social environment. Factors that would determine this relative cognitive impact would include the social(ization) interaction time of the parent with its child, as well as the effort and devotion that the parent spends to socialize its child to the chosen DPI.[6]

We now obtain the formation of the PI that a child $\tilde{a} \in \tilde{A}$ adopts through the 'direct vertical and oblique socialization' process, $\phi_{\tilde{a}}$, as

$$\phi_{\tilde{a}} = \hat{\sigma}_a \phi_a^d + (1 - \hat{\sigma}_a) \phi_{A_a}^d. \tag{1}$$

We will call this the *parental socialization technique*. It embodies the view that the parents set a PI benchmark, ϕ_a^d, and can invest into their parental socialization success share, $\hat{\sigma}_a$, to countervail the socialization influence that the child is exposed to in its general social environment, $\phi_{A_a}^d$. Since the final adopted PI of a child is by construction a convex combination of all DPIs

[5] This terminology stems from Cavalli-Sforza and Feldman [9], and is distinguished from 'horizontal socialization', viz. the socialization influence of members of the same generation (which we leave unconsidered in the present paper).

[6] See e.g. Grusec [12] for an introductory overview of theories on determinants of parental socialization success.

that it observes, the set of possible PIs that a child can adopt then coincides with the convex hull of the set of possible DPIs, con $\phi^d(X) \subseteq \mathbb{R}$.

We assume next that, in their adult life period, all individuals keep the PI that they have adopted in their childhood in an unchanged way. These adopted PIs of the adults can be interpreted to induce 'filters' under which adults can compare and rank different choices of socio–economic actions. This form of evaluation takes place in terms of comparing the DPIs of the socio–economic actions to the own adopted PIs.[7] Specifically, we assume that the adopted PIs induce complete and transitive preference relations over choices of DPIs (respectively the underlying socio–economic actions).

Assumption 1 ('Own' Utility)**.** *For every $a \in A$,*

(a) there is an 'own' utility function $u^{\phi_a} : \text{con } \phi^d(X) \mapsto \mathbb{R}$, $u^{\phi_a}\left(\phi^d_a\right) \in \mathbb{R}$, *where*

(b) u^{ϕ_a} is single–peaked with peak ϕ_a, thus strictly increasing/decreasing at all $\phi^d_a \in \text{con } \phi^d(X)$ such that $\phi^d_a < / > \phi_a$.

Intuitively, the single–peakedness property means that we assume adults to prefer choosing behaviors (DPIs) that are as close as possible in line with their adopted PIs.

2.2 Endogenous Cultural Formation of Preferences

In the present section, we will lay down one specific way of achieving an endogeneization of the cultural formation of preferences process. This will be based on purposeful socialization decisions of parents. Thereby, we notably restrict the latter to consist of their choice of a displayed preference intensity and of their parental socialization success share. This means that we leave the oblique socialization weights (that determine the children's relative social learning from the different adult subsets) exogenously fixed.

Motivation for Purposeful Socialization In a first step, we have to clarify what motivation parents have to actively engage in their children's socialization process, i.e. what induces them to purposefully employ their socialization technique (the functioning of which we assume them to be fully aware of). Basically, we let this motivation stem from the fact that parents also obtain an inter–generational utility component. Thereby, this is either related to the adopted PI of their children and/or to the DPI (respectively the underlying socio–economic actions) that they expect their adult children to take.

As far as the latter expectations are concerned, we make here an assumption on a specific form of parental myopia: Although parents obtain

[7] This is in line with the cognitive dissonance theory of Festinger [11].

an inter–generational utility component, which eventually induces them to choose a DPI that does not coincide with their adopted PI (see below), we assume that they do not realize that this form of behavior changing impact will also be present in their adult children's decision problems. Thus, any parent $a \in A$ expects its adult child to choose a DPI that is in the set of maximizers of its 'own' utility function, $\arg\max_{\phi^d_{\tilde{a}} \in \phi^d(X)} u^{\phi_{\tilde{a}}}\left(\phi^d_{\tilde{a}}\right)$. Under the following assumption, $\phi^d(X)$ is convex and thus $\phi^d(X) = \text{con } \phi^d(X)$. This then guarantees by the single–peakedness of the utility functions that $\arg\max_{\phi^d_{\tilde{a}} \in \phi^d(X)} u^{\phi_{\tilde{a}}}\left(\phi^d_{\tilde{a}}\right) = \phi_{\tilde{a}}$, $\forall a \in A$. Hence, the parental expectations of their adult children's DPIs are uniquely determined.

Assumption 2 (Convexity)**.** *$X \subseteq \mathbb{R}^n$ is non–empty and convex and ϕ^d is continuous. If $n > 1$ then ϕ^d is additionally concave.*

Given the parents' myopic expectations, it is independent of whether the inter–generational utility component of a parent is related to the adopted PI or expected DPI of its adult child, since they coincide. Under this property, we will now assume that any parent perceives an *optimal preference intensity*, such that if the adult child adopts this optimal PI (and is expected to behave according to it), then this is considered by the parent to be 'inter–generational utility maximal'. These parent–specific optimal PIs are subject to what we call *perception rules*.

Thereby, the perception rule of the optimal PI of any parent is determined by two 'ingredients'. The first one specifies a (set of) subset(s) of adults, which can be understood as reference group(s). The second ingredient then specifies the construction of the optimal PI that a parent perceives out of characteristics of the adults in these reference group(s). These characteristics can either be observable (notably the DPIs of adults) or known to an individual parent.

To formally introduce the concept of perception rules, it will be convenient to define $\mathcal{A}$ as a σ–algebra generated by the finite partition $\{A_J\}_{J=1}^K$ (this is without further loss of generality).

Definition 1 (Perception Rule)**.** *The rule for the perception of the optimal PI of parent $a \in A$ is a pair $\left(R_a, \hat{\phi}_{\tilde{a}}\right)$, where $\emptyset \neq R_a \subseteq \{a\} \cup \mathcal{A}$ and where $\hat{\phi}_{\tilde{a}} : \{a\} \cup \mathcal{A} \mapsto \text{con } \phi^d(X)$, $\hat{\phi}_{\tilde{a}}\left(R_a\right) \in \text{con } \phi^d(X)$.*

To ease the interpretation of this conceptualization, we will list here three sensible types of perception rules for optimal PIs. In section 3, we will, in a two cultural groups setting, be concerned with analyzing evolutionary processes subject to the second and third type of perception rules mentioned here.[8] Note also that the list below is not meant to be exhaustive (one could e.g. consider combinations of the three types mentioned).

[8]The first, 'imperfect empathy', type has already been discussed in Pichler [16]. In Appendix B.1, the respective results for the two cultural groups setting are shortly discussed.

CR 1 The optimal PI of a parent $a \in A$ is identical to its adopted PI, $R_a = \{a\}$ and $\hat{\phi}_{\tilde{a}}(\{a\}) = \phi_a \in \text{con } \phi^d(X)$.

CR 2 The optimal PI of a parent $a \in A$ is identical to a parent–specific (model–exogenous) PI, $R_a = \{a\}$ and $\hat{\phi}_{\tilde{a}}(\{a\}) = e_a \in \text{con } \phi^d(X)$.

CR 3 The optimal PI of a parent $a \in A$ is identical to the average DPI of one of the adult subsets, $R_a = A_M$, $M \in \{1, \ldots, K\}$, and $\hat{\phi}_{\tilde{a}}(A_M) = \phi^d_{A_M} \in \text{con } \phi^d(X)$.

Given the perception rule rules and the resulting optimal PIs, we assume further that parents perceive utility losses for deviations of the adopted PI of their children from these optimal PIs. Specifically, for any parent $a \in A$, we introduce the parameter $i_a \in \mathbb{R}_+$ that shall capture the strength of the perceived inter–generational utility losses. We will call this the parent's *inter–generational preference intensity.* For simplicity, we assume that these are invariably passed over from an adult to its child, $i_{\tilde{a}} = i_a$, $\forall a \in A$.

Assumption 3 (Inter–generational Utility)**.** *$\forall a \in A$,*

(a) there is an inter–generational utility function $v^{\hat{\phi}_{\tilde{a}}(R_a)}(\cdot\,|i_a) : \text{con } \phi^d(X) \mapsto \mathbb{R}$, $v^{\hat{\phi}_{\tilde{a}}(R_a)}(\phi_{\tilde{a}}\,|i_a) \in \mathbb{R}$, where

(b) $\forall i_a \in \mathbb{R}_{++}$, $v^{\hat{\phi}_{\tilde{a}}(R_a)}(\cdot\,|i_a)$ is single–peaked with peak $\hat{\phi}_{\tilde{a}}(R_a)$, thus strictly increasing/decreasing at all $\phi_{\tilde{a}} \in \text{con } \phi^d(X)$ such that $\phi_{\tilde{a}} < / > \hat{\phi}_{\tilde{a}}$.

Optimization Problems and Nash Equilibrium In the last step toward the construction of the parental optimization problems, let us finally discuss the cost associated with investments into controlling the parental socialization success share. These would concern e.g. the opportunity cost of the time parents spend for the active socialization of a child, as well as the (psychological) cost of the effort and devotion invested. We will represent these cost by an indirect cost function of choices of socialization success shares. This function is assumed to be identical for all adults $a \in A$ and will be denoted $c : [0,1] \mapsto \mathbb{R}_+$, $c(\hat{\sigma}_a) \in \mathbb{R}_+$.

The parental optimization problem is it then to choose a DPI and its socialization success share such as to maximize the life–time utility net of the cost of achieving the chosen socialization success share. Assuming (for analytical simplicity) additive separability of the utility and cost functions, we obtain, $\forall a \in A$,

$$\max_{\left(\phi^d_a, \hat{\sigma}_a\right) \in \phi^d(X) \times [0,1]} u^{\phi_a}\left(\phi^d_a\right) + v^{\hat{\phi}_{\tilde{a}}(R_a)}(\phi_{\tilde{a}}\,|i_a) - c(\hat{\sigma}_a) \tag{2}$$
$$\text{s.t. } \phi_{\tilde{a}} = \hat{\sigma}_a \phi^d_a + (1 - \hat{\sigma}_a)\phi^d_{A_a}.$$

The optimization problems of the parents hence basically consist of trading off the cost and benefits of their socialization choices. The cost are constituted by 'own' utility losses that parents experience when choosing a DPI that deviates from their adopted PI, together with the cost of a choice of their socialization success share. The benefits accrue in form of resulting inter–generational utility gains through reductions in the distance between the child's adopted PI and the perceived optimal PI.

Finally, let us briefly consider issues on the existence of a (pure strategy) Nash equilibrium of the game that is induced by the strategic interdependence of the individual parental choices. These can take the following forms.

First of all, as has already been discussed, the net life–time utility of an individual parent, i.e. the object of its optimization problem (2), depends on the location of the representative DPI of the general social environment. This is constructed out of the oblique socialization weights and the average DPIs of the adult subsets. Second, the decisions of the other adults could influence the net life–time utility of an individual parent via the perception rule for its optimal PI (as e.g. in the third type of perception rule).

In this respect, to guarantee the existence of a Nash equilibrium, the following additional normalization is required: If the perception rule of a parent is based on the DPIs and/or socialization success shares of one or more subsets of the adults, then this is only in terms of the respective averages.

Proposition 1 (Nash Equilibrium Existence[9]). *If Assumptions 1–3 hold, and if X is compact and the functions $\hat{\phi}_{\tilde{a}}$ are continuous for every $a \in A$, then a Nash equilibrium exists.*

Proof. Confer Proposition 3 in Pichler [16].

The existence result above means that in any given period, we can use (a selection of) the Nash equilibrium choices for substitution in the formation of PIs equation (1). By doing so, we obtain an endogenous representation of the inter–generational formation of PIs, i.e. we have endogenized the cultural formation of preferences process.

3 Assimilation of Cultural Groups

In this section, we will embed the endogenous cultural formation of preferences framework in an environment where the society is populated by two distinct cultural groups. The focus of the subsequent subsections will be on the analysis of the evolution of the adopted preference intensities and induced preferences subject to the Nash equilibrium socialization decisions

[9] The concept of a Nash equilibrium that we employ here requires that almost all parents play best replies. This follows Schmeidler [18] and Rath [17].

of the parents of both cultural groups. This will be done by imposing two distinct types of perception rules. In subsection 3.1 we will consider the second type of perception rule discussed above, while as section 3.2 is based on the third type. Finally, the results for the first, 'imperfect empathy', type of perception rule in the present setting are shortly discussed in Appendix B.1.

Consider any given period $t \in \{0, 1, \dots\}$ and assume that the adult set of that period is partitioned in two subsets with strictly positive measure, $A_t = H_t \cup L_t$. We will call these subsets *cultural groups*, and index them $G_t \in \{L_t, H_t\}$, with population shares $q_{G_t} := \lambda(G_t)$. In the present setting, it will be convenient to index the members of the adult generations in any period t as $g_t \in G_t$, and to denote $-G_t := \{L_t, H_t\} \backslash \{G_t\}$.

Assume now that in the given period, the *parent–child profiles* are group–symmetric. With this we mean that within a cultural group, all adults have identical adopted PIs, which we denote ϕ_{G_t}; identical inter–temporarily fixed inter–generational PIs, denoted i_G; and identical inter–temporarily fixed perception rules for optimal PIs, denoted $\left(R_G, \hat{\phi}_G\right)$. Finally, we assume the oblique socialization weights of *all* children to be also identical. As far as the latter are concerned, we even assume *unbiased* oblique socialization. In this case, the representative DPIs are identical for all children of the society and coincide with the average DPI of the adults. We will denote this average DPI $\phi^d_{A_t}$.

Under these normalizations, we obtain the optimization problems (2) of any parent $g_t \in G_t$, $G_t \in \{L_t, H_t\}$ as

$$\max_{\left(\phi^d_{g_t}, \hat{\sigma}_{g_t}\right) \in \, \phi^d(X) \times [0,1]} u^{\phi_{G_t}} \left(\phi^d_{g_t}\right) + v^{\hat{\phi}_G(R_G)} \left(\phi_{g_{t+1}} \left| i_G\right.\right) - c\left(\hat{\sigma}_{g_t}\right) \tag{3}$$
$$\text{s.t. } \phi_{g_{t+1}} = \phi^d_{g_t} + (1 - \hat{\sigma}_{g_t})(\phi^d_{A_t} - \phi^d_{g_t}).$$

Given the group–symmetry that we have established, it follows that all parents of a cultural group do always have identical sets of pairs of best reply choices of a DPI and socialization success share. In a further extension of the symmetry–properties, we will assume at this point that even all parents of a cultural groups always select the same best reply pairs out of these sets (but this is not known to the single individual).[10]

Assumption 4 (Compactness, Continuity)**.** *X is compact, and the functions* u^b, $v^d\left(\cdot\, |h\right)$, *c and* $\hat{\phi}_G$ *are continuous.*

Subsequently, we will call a *symmetric Nash equilibrium (SNE)* a Nash equilibrium where all parents of a cultural group that choose a best reply do also choose the same one.

[10]Alternatively, this property would be satisfied if we would assume the target functions of the optimization problems of all parents to be strictly concave (since then the best reply sets would be single–valued).

Corollary 1 (Symmetric Nash Equilibrium)**.** *Under the group–symmetry properties and Assumption 4, a* symmetric Nash equilibrium (SNE) *exists in any period.*

Proof. The existence has been discussed in Proposition 1, and the symmetry holds basically by assumption. □

Within the set–up of the present section, the set of SNEs of any period depends on the adopted and inter–generational PIs, the perception rules, as well as on the population shares of the two cultural groups. The subsequent analysis will be based on the selection of a single element of these SNE sets for any $P_t := \{\phi_{L_t}, \phi_{H_t}, i_L, i_H, \left(R_L, \hat{\phi}_L\right), \left(R_H, \hat{\phi}_H\right), q_{H_t}\} \in \phi^d(X)^2 \times \mathbb{R}_+^2 \times \left(\mathcal{A} \times \mathcal{C}^0\right)^2 \times [0,1]$. We will denote the selected SNE as a tuple

$$\left\{\phi^{d^*}_{G_t}(P_t), \hat{\sigma}^*_{G_t}(P_t)\right\}_{G_t=L_t,H_t}.$$

Substituting these into the parental socialization techniques (1), it follows that almost all children of both cultural groups adopt the same PIs. We obtain these as

$$\phi_{G_{t+1}} = \phi^{d^*}_{G_t}(P_t) - \left(\phi^{d^*}_{G_t}(P_t) - \phi^{d^*}_{-G_t}(P_t)\right)\left(1 - \hat{\sigma}^*_{G_t}(P_t)\right)(1 - q_{G_t}). \tag{4}$$

This constitutes a representation of the members of the two cultural groups that is inter–generationally 'continuous'. With this, we mean that (almost all) contemporaneous children of a cultural group form next period's adults of the same cultural group. It then also follows that the population shares of the two cultural groups are constant and we will drop the respective time–indexes subsequently.

Integration and Assimilation The analysis in the succeeding two subsections will always be initiated by a discussion of the SNE choices of any given period under the different types of perception rules. In this context, we will speak of *behavioral dis–integration* of the adult members of a cultural group $G_t \in \{L_t, H_t\}$ whenever it holds that $\left|\phi^{d^*}_{G_t}(P_t) - \phi^{d^*}_{-G_t}(P_t)\right| > \left|\phi_{G_t} - \phi^{d^*}_{-G_t}(P_t)\right|$. This means that these adults choose a more 'radical' DPI relative to the DPI of the other group's adults than the choice of their adopted PI would mean.

In an inter–temporal context, it will be crucial to determine the endogenous evolution of the SNE choices — and with it the endogenous evolution of the adopted PIs. Specifically, we will also want to answer the question of the inter–temporal assimilation (or dissimilation) process between the two cultural groups. In a slight variation of the terminology introduced in Pichler [16], we will speak of *(PI) assimilation* whenever the PI–distance

$\Delta_t^{\phi} := |\phi_{L_t} - \phi_{H_t}|$ strictly declines over generations, i.e. $\Delta_{t+1}^{\phi} < \Delta_t^{\phi}$. From equation (4), we obtain the PI–distances under SNE choices as

$$\Delta_{t+1}^{\phi} = \left|\left(\phi_{L_t}^{d^*}(P_t) - \phi_{H_t}^{d^*}(P_t)\right)\right| \left(\hat{\sigma}_{L_t}^{*}(P_t) q_H + \hat{\sigma}_{H_t}^{*}(P_t)(1 - q_H)\right). \tag{5}$$

Furthermore, if the assimilation is such that the adopted PIs of the members of the cultural group with the contemporaneously smaller PI strictly increase over generations, while the opposite holds vice verso, we will speak of *strict assimilation*.

Finally, with *behavioral assimilation*, we will call a situation where the absolute distance between the SNE choices of DPIs of the two groups strictly declines between two generations.

3.1 Fixed Optimal Preference Intensities

In the present subsection, we consider a situation where the parents of both cultural groups perceive (exogenously given) inter–generationally fixed optimal preference intensities. Thus, in any given period and for both $G_t \in \{L_t, H_t\}$, $\hat{\phi}_G(R_G) = e_G \in \operatorname{con}\phi^d(X)$. This structure corresponds to the second type of perception rule. Without loss of generality, consider subsequently the (non–degenerate) case where $e_H > e_L$.

The following assumption will be prerequisite for a meaningful characterization of the symmetric Nash equilibrium choices.

Assumption 5 (Slope)**.**

(a) u^b *and* $v^d(\cdot|h)$ *are differentiable at their peaks, and*

(b) c *is differentiable at the origin with slope zero, and strictly increasing in the interval* $(0, 1]$.

Since both the utility and inter–generational utility function are single peaked, it follows by Assumption 5 (a) that both functions have zero slope at their peaks. Thus, parents perceive no (inter–generational) utility losses for marginal deviations of their chosen DPI from their adopted PI, respectively of their adult child's adopted PI from the optimal PI.

In the rest of the analytical part of this subsection, we will be concerned with characterizing the SNE choices of the parents as well as the resulting evolutions of the PIs of the two cultural groups. To do this, we will focus our attention on what we call the *generic state space*.

Proposition 2 (Generic State Space)**.** *Let Assumptions 1–5 hold. Then,* $\forall P_0 \in \phi^d(X)^2 \times \mathbb{R}_{++}^2 \times \left(\mathcal{A} \times \mathcal{C}^0\right)^2 \times (0,1)$ *subject to* $\hat{\phi}_H(R_H) = e_H > e_L = \hat{\phi}_L(R_L)$*,* $\exists \infty > T(P_0) \geq 0$ *such that* $e_H > \phi_{H_{T(P_0)}} > \phi_{L_{T(P_0)}} > e_L$.

Proof. In Appendix A.1.

This latter proposition states the following. Independent of the initial PIs of the two cultural groups, the PIs will enter a basin in the state space where the positions of the two PIs can be considered 'consistent' with the relative location of the fixed optimal PIs. With this we mean that (a) the group with the strictly larger fixed optimal PI does also have a strictly larger adopted PI, and (b) the PIs of both groups do lie in the interior of the 'PI–space' that is formed by the two fixed optimal PIs.

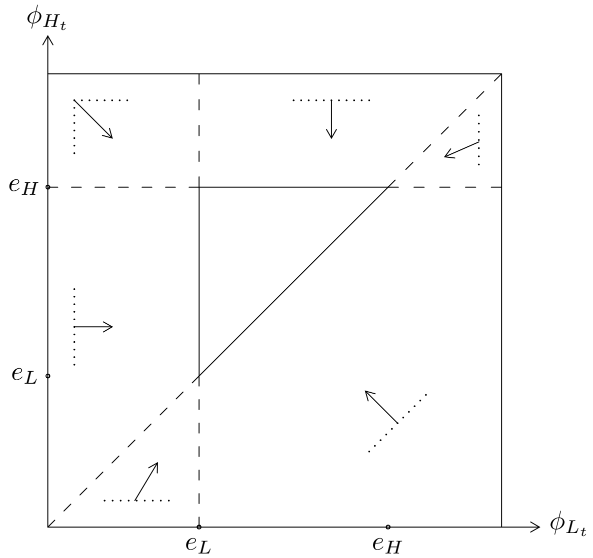

Figure 1: Phase Diagram (Non–generic State Space)

That any path that starts outside this generic state space must lead into it is illustrated in the phase diagram 1. In any of the fields in this diagram, the dotted lines indicate the boundaries of the range of the angles that the phase vectors can take (notably, the boundaries themselves are not included in this range). Also, one of these possible phase vectors is always depicted. Furthermore, the phase vectors on the boundaries between the various fields share (the combination of) the properties of those in their neighboring fields. This also implies that all phase vectors on the boundary of the generic state space point into it.

Let us briefly discuss the basic intuition to understand this phase diagram. We start with the two ('non–generic') fields in the upper triangle of the state space where the PI of group L is smaller than optimal. This implies that the direction of the socialization efforts of the members of this group is 'upwards' (i.e. they tend to choose a DPI that is larger than their adopted PI, jointly with a strictly positive parental socialization success share). Since also both the adopted PI and the optimal PI of group H are strictly larger than the adopted PI of group L, their chosen DPI tends to be strictly larger than the adopted PI of group L. This combination leads to a strict inter–generational increase of the adopted PI of group L under SNE choices. The analogous logic shows that, within the fields in the upper triangle of the state space where the PI of group H is larger than optimal, the adopted PI of group H must strictly decrease.

Consider now the lower left triangle in the state space. In this, the PI of group H is smaller than that of group L, and both are smaller than the optimal PI of group L. In this case, the directional socialization efforts of both groups are (strictly) 'upwards'. This implies that at least the adopted PI of group H must strictly increase inter–generationally. Again, the analogous logic shows that in the upper right triangle, the adopted PI of group L must strictly decrease.

Finally, consider the lower right field in the state space. In this, the PI of the members of cultural group L is larger than both their adopted and optimal PI, as well as larger than the adopted PI of cultural group H. Furthermore, the latter is smaller than optimal. This implies that the directional socialization effort of the members of group L is 'downwards' while that of group H is 'upwards'. This combination then yields the effect that under SNE choices, the inter–generational increase in the adopted PI of group H must be strictly larger (respectively strictly less negative) than that of group L.

We will now turn to the characterization of SNE choices within the generic state space. Note that the results below do hold for all elements in the sets of SNE choices of any period (thus, for any SNE selection function).

Proposition 3 (SNE Characterization). *Let Assumptions 1–5 hold and let* $e_H > \phi_{H_t} > \phi_{L_t} > e_H$. *Then,* $\forall\{i_L, i_H, q_H\} \in \mathbb{R}^2_{++} \times (0,1)$,

(a) $\phi^{d*}_{H_t}(P_t) > \phi_{H_t} > \phi_{L_t} > \phi^{d*}_{L_t}(P_t)$,

(b) $\hat{\sigma}^*_{G_t}(P_t) \in (0,1]$, $\forall G_t \in \{L_t, H_t\}$,

(c) $e_H > \phi_{H_{t+1}} > \phi_{L_{t+1}} > e_L$.

Proof. In Appendix A.2.

Within the generic state space, the socialization efforts of the members of the two cultural groups are in the opposite directions. This yields the result that in any SNE, the parents of both cultural groups dis–integrate behaviorally and choose strictly positive socialization success shares. Nevertheless, their socialization investments would never be intense enough such that the next generation's adopted PIs would exactly coincide with the optimal one (the logic of this sort of result is being discussed in Pichler [16]). This means that once the PIs of the two groups have entered the generic state space, they will never leave it again. Thus, in an extension of Proposition 2, it follows that for every $t' \geq T(P_0)$, $e_H > \phi_{H_{t'}} > \phi_{L_{t'}} > e_L$.

Obviously, also any steady state must be located within the generic state space. The following assumption is prerequisite for guaranteeing asymptotic stability of at least one such rest point.

Assumption 6 (Continuous SNE Selection Function)**.** *A continuous SNE selection function exists.*

Proposition 4 (Stable Steady State)**.** *Let Assumptions 1–6 hold. Then, an asymptotically stable steady state exists (in the generic state space).*

Proof. In Appendix A.3.

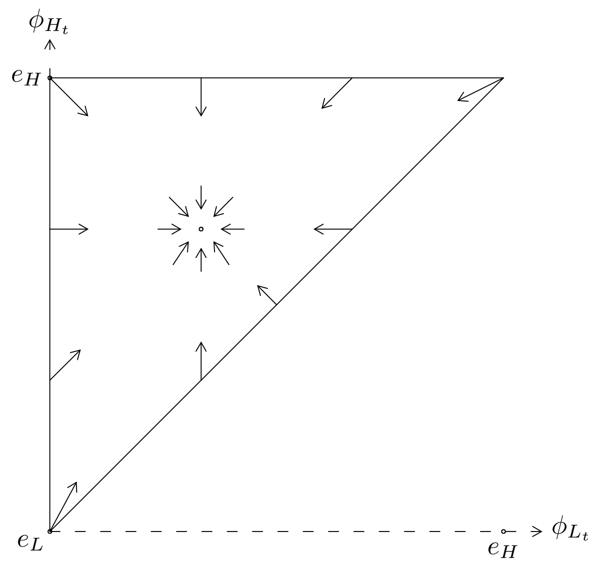

Figure 2: Phase Diagram (Generic State Space)

This stability result is driven by the combination of the following forces. First, as has been shown in Proposition 3, within the generic state space, the direction of the socialization efforts of the members of both cultural groups are toward their fixed optimal PIs — thus in opposite directions. This also implies that both types of parents have to countervail the unfavorable preference influence that the DPI of the members of the other cultural group have on their children. These countervailing efforts tend to be increasing for increasing distances between the adopted PIs of the two groups. Second, the intensity of the socialization efforts of the parents of both cultural groups tend to be decreasing with decreasing distance of their adopted PIs to their fixed optimal PI (the latter two properties are shown in the comparative statics Appendix B.2).

Within a neighborhood of the boundary of the generic state space, the combination of those effects is always unbalanced, i.e. no steady state exists therein. However, a rest point must exist in the generic state space. Even, there is one such point for which it holds that in a neighborhood around it, the above described effects are unbalanced in a way that guarantees asymptotic stability.

The phase diagram 2 illustrates the just described dynamic forces within the generic state space.

Numerical Illustration We will close this subsection with an illustration of the analytical findings by means of a numerical simulation.[11] In both cases of Figure 3, the initial PIs of the members of both groups coincide with their fixed optimal PIs. Furthermore, group L is always the minority with a population share of ten per cent. Such a constellation can e.g. be interpreted as having resulted from immigration (of group L), where initially the hosting and immigrant group have had adopted exactly their optimally perceived PI. The latter could e.g. be constantly indoctrinated by two distinct religious institutions which the two different cultural groups adhere to — and which thus constitute the respective norms on behavior of the two groups.

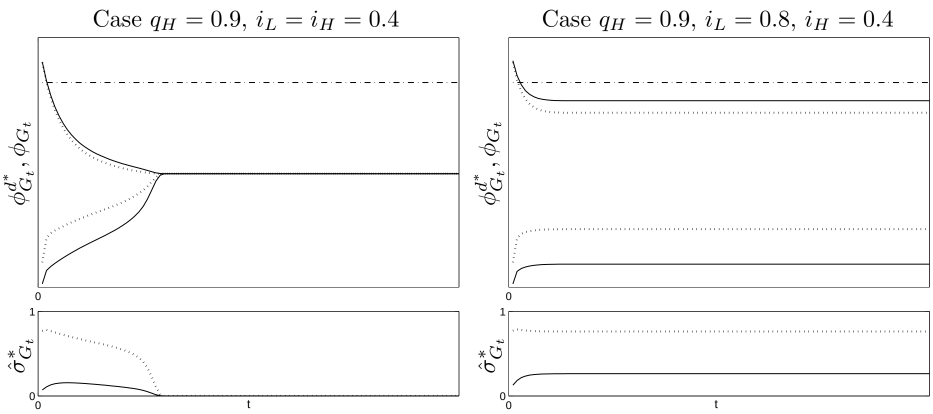

Figure 3: Evolution under Fixed Optimal PIs

In the upper graph of each case, the solid and dotted lines represent

[11]For the numerical illustrations in this and the following subsections, we use the following specifications (unless otherwise noted): $u^b(j) = -(b-j)^2$, $v^d(k|h) = -h(d-k)^2$, $c(o) = o^2$.

Furthermore, in the present cases, we used an initial PI–distance of 4 units (which thus coincides with the distance between the fixed optimal PIs), and the total length of the time–axis in all graphs corresponds to 100 periods. Nevertheless, the choice of the initial PI–distance and the resulting length of the convergence path could be arbitrarily changed, so that they are not indicated in the graphs.

the DPIs and PIs of the two groups (and the dash–dotted line locates the population–share weighted convex combination of the initial PIs; this would equate to the steady state if parents of both cultural groups would not invest into their socialization instruments). The paths of the socialization success shares of the parents of group L are represented by the dotted lines in the lower graph of each case.

The left case of the figure stylizes immigration of a cultural group with a comparatively weak norm on behavior into a hosting society for which this also holds. To the contrary, the right case of the figure stylizes immigration of a cultural group with a strong norm on behavior into a hosting society with a comparatively weak norm.

Let us first collect the evolutionary regularities that can be seen in both cases. First, the members of both cultural groups dis–integrate behaviorally in every period. Second, there is an assimilative tendency until the steady state has been reached. This PI–assimilation is accompanied by an assimilation of the chosen DPIs of the cultural groups (see also the comparative statics results in Appendix B.2).

Furthermore, in both cases the minority cultural group invests considerably more into both socialization instruments. This is (partially) due to the fact that the minority group faces a much more unfavorable composition of the general social(ization) environment (in terms of the resulting location of the average DPI compared to the fixed optimal PI). The individual parents of that group thus aim to compensate this by increased investments into their socialization instruments. In the right case, the minority group does additionally have a much stronger norm on behavior, i.e. the social punishments from behavioral deviations from the norm are accordingly more intense. This additionally induces the parents of this group to invest more into socialization.

This latter effect has a remarkable impact on the dynamical evolution of the endogenous variables. In the left case, the norms of both cultural groups are low enough such as to allow for a substantial assimilation process. Even, the PIs of the two groups do nearly converge to a symmetric steady state (but stay distinct).[12] Compared to this case, the increased socialization investments of the minority group in the right case do trigger an according reaction of the majority group. Thus, both groups invest more into both socialization instruments. As a consequence, the PIs of the two groups are held back from assimilation already after very small deviations from the fixed optimal PIs. The resulting steady state PI–distance is accordingly larger.

[12] We call a *symmetric steady state* a steady state where almost all adults have the same adopted PI.

3.2 Endogenous Norms on Behavior

The present section will be based on the third type of perception rule discussed in section 2.2. The latter lets all parents form their perception of the optimal preference intensity based on a the average DPI of a subset of the adults. In the present context, it is immediate to let the respective subsets coincide with the adult members of the own cultural group of a parent. Thus, under the symmetry assumptions taken above, the optimal PI always coincides with the identical best reply DPIs of the own group members.

We will start the analytical part of this subsection with a characterization of the SNE choices of any period. To be able to do this in a sensible and consistent way, we will require the additional assumption below.

Assumption 7 (Concavity)**.** *The target functions of optimization problems (3) are concave.*

This assumption is stronger than it might appear on first glance. To see this note that concavity of the own and inter–generational utility functions together with convexity of the cost function is not in general sufficient to guarantee concavity of the target functions of the optimization problems. This follows since the Hessian matrices of the parental socialization techniques with respect to the two decision variables are indefinite.[13] Thus the inter–generational utility functions are not in general concave with respect to the two decision variables. To cure this, it is thus necessary that the own utility functions together with the cost functions are jointly concave and convex enough compared to the concavity of the inter–generational utility functions.

Once this sort of condition is satisfied, it is then also guaranteed that in any given period, SNE choices exist that preserve the relative positions of the adopted PIs of the two groups. To require this property can be considered sensible, since it assures a minimum sort of continuity of the inter–generational evolution of the PIs.

Proposition 5 (Relative–Position–Preserving SNE Choices)**.** *Let Assumption 1–5 and 7 hold. Then, in any given period and for every* $\{i_L, i_H, q_H\} \in (\mathbb{R}_{++} \setminus \{\infty\})^2 \times (0,1)$*, a SNE with the following characteristics exists.*

1. *Case* $\phi_{H_t} >< \phi_{L_t}$
   - *(a)* $\phi^{d*}_{H_t}(P_t) >< \phi_{H_t} >< \phi_{L_t} >< \phi^{d*}_{L_t}(P_t)$,[14]
   - *(b)* $\hat{\sigma}^*_{G_t}(P_t) \in (0,1)$, $\forall G_t \in \{L_t, H_t\}$,
   - *(c)* $\phi^{d*}_{H_t}(P_t) >< \phi_{H_{t+1}} >< \phi_{L_{t+1}} >< \phi^{d*}_{L_t}(P_t)$.

[13] The determinants of these Hessian matrices are -1.

[14] The outer inequalities turn into equalities if the adopted PI of a cultural groups equals the relevant one of the boundaries of the set of possible DPIs.

2. *Case* $\phi_{H_t} = \phi_{L_t}$

   (a) $\phi^{d*}_{G_t}(P_t) = \phi_{G_t} = \phi_{G_{t+1}}$, $\forall G_t \in \{L_t, H_t\}$,

   (b) $\hat{\sigma}^{*}_{G_t}(P_t) = 0$, $\forall G_t \in \{L_t, H_t\}$.

Proof. In Appendix A.4.

The key to understanding these properties is the following. First remember that under the present perception rule, the optimal PI of the members of both cultural groups always coincides with their identical best reply DPI choices. This opens the possibility of multiple, qualitatively different, SNEs. However, consider a situation where the DPIs are such that both groups dis–integrate behaviorally. Then, the best reply directions of socialization efforts would coincide with this constellation, i.e. there would be best reply behavioral dis–integration of both groups. This is the basis for the existence of a SNE that as characterized in the first part of Proposition 5.

Since both cultural groups dis–integrate behaviorally, together with a strictly positive socialization success share, it also follows that the relative PI positions of the two groups are preserved over generations. However, the parents of both cultural groups would never choose to exclusively socialize their children (choose a parental socialization success share of one). This follows since in this case, their adult children's adopted PI would coincide with the chosen DPI of the parents, thus with the optimal PI. This can though never be subject to best reply choices (as discussed in detail in Pichler [16]).

Finally, in the case where the adopted PIs of both cultural groups are identical, the situation where the parents of both cultural groups do not actively socialize their children is possible under SNE choices. This follows since such a choice–constellation would yield maximum possible utility for all parties involved. Notably, since the adopted PIs of all adult children would then coincide with the adopted PIs of the contemporaneous adult generation, any such case would constitute a steady state (which is additionally relative position preserving).

Under these relative position preserving properties, it furthermore follows that no PI–trajectory that has its origin in the upper/lower triangle of the state space can enter the lower/upper triangle. We will next turn to the discussion of the qualitative properties of the corresponding PI–dynamics. For one part of this discussion, the following assumption will be required.

Assumption 8 (Symmetric Utility Functions). *For every* $b, b' \in \operatorname{con} \phi^d(X)$, $u^b(j) = u^{b'}(j')$ *if* $b - j = b' - j'$. *Similarly, for every* $d, d' \in \operatorname{con} \phi^d(X)$, *and* $h \in \mathbb{R}_+$, $v^d(k|h) = v^{d'}(k'|h)$ *if* $d - k = d' - k'$.

This assumption states that all 'own' and inter–generational utility functions yield identical felicity for identical 'directional' deviations from their peaks. On the one hand, this assumption appears to be quite natural. On the other hand, note that it implies that the dis–utilities that accrue due to deviations from the utility peaks are independent of the positions of the peaks relative to the boundaries of the set of possible DPIs (respectively 'adoptable' PIs).[15]

Proposition 6 (Evolution under Endogenous Norms on Behavior)**.** *Let Assumptions 1–7 be satisfied. Then, the following properties are satisfied under relative position preserving SNE choices and for every* $(i_L, i_H, q_H) \in \mathbb{R}_+^2 \times (0,1)$*.*

(a) There exists a $\Delta(i_L, i_H, q_H) \in \big(0, |\operatorname{con}\phi^d(X)|\big]$ *such that* $\forall 0 < \Delta_t^{\phi} < \Delta(i_L, i_H, q_H)$*,* $\Delta_{t+1}^{\phi} < \Delta_t^{\phi}$*. This implies that* $\forall \Delta_0^{\phi} < \Delta(i_L, i_H, q_H)$*,* $\lim_{t\to\infty} \Delta^{\phi}(t, \Delta_0^{\phi}, i_L, i_H, q_H) = 0$*.*

(b) If additionally Assumption 8 is satisfied, then there exists a SNE selection function such that $\forall (\phi_{L_0}, \phi_{H_0}) \in \operatorname{con}\phi^d(X)^2$*, the PIs converge to a steady state.*

Proof. In Appendix A.5.

The first part of this proposition states that indeed, there is a basin in terms of a maximum PI–distance such that for any pair of PIs that features a lower distance, the cultural groups assimilate inter–generationally.

This result rests crucially on the continuity of the relative position preserving SNE selection function. Remember that we imposed the normalization that at any symmetric PI point (i.e. any point on the main diagonal of the state space), the parents of both cultural groups choose zero investments into their socialization instruments. Thus, any such a point is a rest point. By the continuity of the SNE choices, it must then hold that in some neighborhood around the main diagonal, the parents of both groups choose low enough pairs of behavioral dis–integration and socialization success shares, such that this results into assimilation. Thus, for any initial pair of PIs of the groups that is located within the latter, the PIs converge to a symmetric steady state.

Nevertheless, the basin of attraction of the symmetric steady states does not in general coincide with the whole state space. The second part of Proposition 6 then states that (under the relevant conditions) even any

[15] To see that accounting for these relative positions might be sensible, consider a pair of unequal adopted PIs. Then, any identical DPI deviation from these utility peaks in the same direction would imply that always one chosen DPI can be considered more 'radical' relative to the maximum or minimum possible DPI. Thus, if one would e.g. like to account for the adults' eventual 'preferences' for moderate behavior, Assumption 8 would not be appropriate. A similar line of thought applies in case that parents would e.g. prefer their adult children having moderate adopted PIs (respectively choosing more moderate DPIs).

initial pair of PIs that is located outside this basin converges. This property rests on normalizations of the phase vectors (which is achieved through additional normalizations of the underlying SNE selection function) to rule out the existence of circles in the whole state space.

In very short words, these normalizations are such that the state space is being composed of a continuum of connected line segments. These consist of (a) a vertical line on which the lower/upper bound of the set of possible DPIs is binding in the DPI choice of the parents of group L (in the upper/lower triangle of the state space), and on which the inter–generational PI–change of group H is constant; (b) a $45\,^{\circ}$–line on which the inter–generational PI–changes of both groups are constant (notably, these lines can 'melt down' to single points); and (c) a horizontal line on which the upper/lower bound of the set of possible DPIs is binding in the DPI choice of the parents of group H (in the upper/lower triangle of the state space), and on which the inter–generational PI–change of group L is constant.

Since the state space is thus constructed as a continuum of (connected) line segments on which the inter–generational PI–change of group L and/or group H are constant, it follows that no circles can exist. Thus, any sequence of PIs must converge.

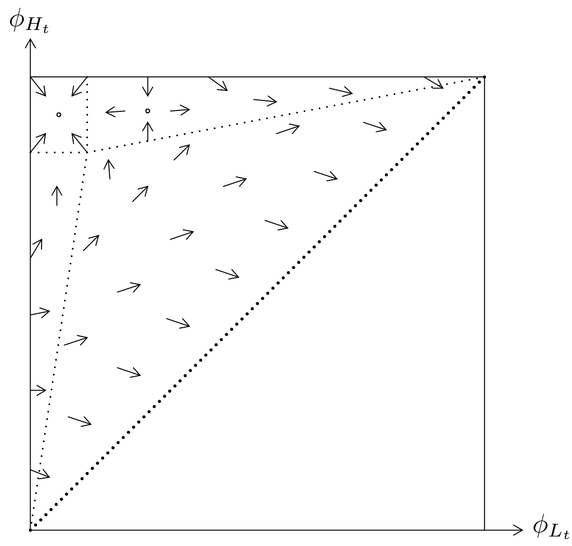

Figure 4: Phase Diagram (Upper Triangle)

The results of Proposition 6 are illustrated in Figure 4, which stylizes possible qualitative properties of the phase vectors in the upper triangle of the state space (the phase diagram in the lower triangle would correspond to the mirror image). This upper triangle is partitioned into four distinct fields, indicated by the dotted lines. These stylize the regions where either the lower/upper bound of the set of possible DPIs is binding for group L/H (the leftmost/upper–rightmost triangle), respectively where both boundaries are binding (the rectangle), or where both boundaries are unbinding (the main triangle).

The central characteristic of this phase diagram is that with increasing PI–distance, the PI–assimilation of the cultural groups declines in magnitude. Specifically, in a neighborhood around the main diagonal (which consists of a continuum of steady states) the cultural groups do first strictly assimilate, followed by a neighborhood in which assimilation takes place. Furthermore, there is a $45\,°$–line in the main triangle where the PI–distance stays constant (but not the PIs themselves in this case).

For any point in the main triangle that features a larger PI–distance, the cultural groups do even (strictly) dissimilate. In the present illustration, where the socialization efforts of the parents of group H are always dominating (which can be due to e.g. a larger strength of the behavioral norm), this has the following consequence: Any PI–trajectory that starts in the according area of the state space must lead into a field where (at least) the upper bound of the set of possible DPIs is binding for group H.

In this field, there is then a separating vertical line with the following properties. If a trajectory enters (or starts in) the field 'to the left' (i.e. at a point with a lower adopted PI of group L) of this vertical line, then the PIs will converge to the asymptotically stable steady state in the rectangle. In the opposite case, the PIs will be subject to an assimilation process toward a symmetric steady state. Finally, if the trajectory should enter the field exactly at the vertical line (or starts thereon), then the depicted unstable steady steady state would be reached.

Numerical Illustration We again conclude this subsection with a numerical illustration of the evolutionary dynamics[16]. In both cases of Figure 5, group L is a again the minority with a population share of twenty per cent. Furthermore, it has a slightly lower intensity of the endogenous norm on behavior.

The only distinction between both cases is that the right case features a twice as high initial PI–distance as the left case. As can be seen, this has a crucial consequence on the evolutionary dynamics. The constellation in the left case is such that the initial PIs are located in the basin of attraction of the symmetric steady states. Even, both cultural groups do assimilate throughout the convergence path. This process is again accompanied by an assimilation of the chosen DPIs.

These results do not hold in the right case. To the contrary, the initial PI

[16]Compared to the previous subsection, the total length of all time–axes is reduced to 30 units.

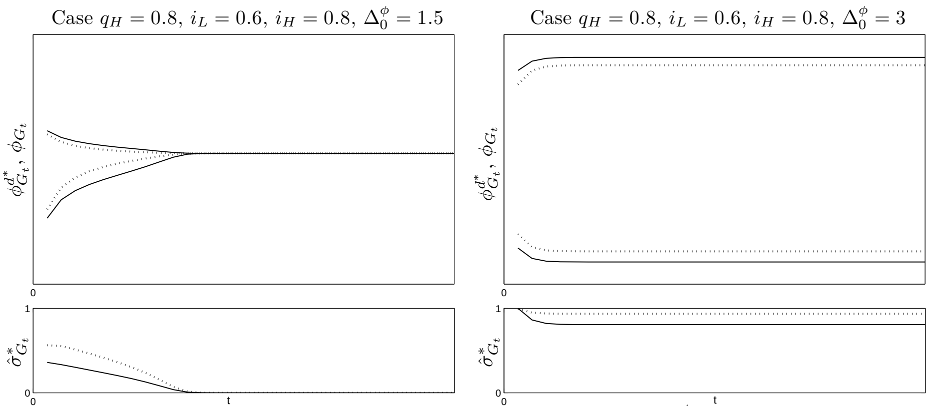

Figure 5: Evolution under Endogenous Behavioral Norms

distance is large enough such as that even an inter–generational dissimilation process is triggered — both with respect to the PIs as well as the chosen DPIs.

4 Conclusions

This paper extended a recent theory of Pichler [16] on the inter–generational formation of continuous preferences. Followed by a recapitulation of the latter theory, we analyzed the dynamic evolution of both the behavior and the preferences in a society populated by two distinct cultural groups.

We showed that the qualitative dynamic properties depend crucially on how parents form their perception of the optimal preference intensity that their children should adopt. As has already been shown in Pichler [16], if all parents have 'imperfect empathy', then the preferences of (almost) all dynasties converge to the same point. To the contrary, if all parents of a cultural group adhere to the same exogenously given and fixed optimal preference intensity, then this can never happen. Rather, the two cultural groups stay distinct forever.

The largest variety of possible qualitative properties of the convergence paths is being featured when the optimal preference intensities of all parents of a cultural group coincide with that derived from the average behavior of the group members. Given this, it is well possible that the preferences of (almost) all parents converge to the same point. However, it can occur that the cultural groups initially assimilate, but stay distinct in the long run. Even, an inter–generational dissimilation process that leads to a steady state with larger than initial preference distance can realize.

This sort of analysis also yields additional insights into empirically observable patterns of assimilation and integration of cultural groups. However, the present one–dimensional framework can only be considered the first step in a longer road toward a holistic representation of these processes. The next steps on this road could concern a general, n–dimensional analysis, both with respect to the number of preference types, as well as to the number of cultural groups. Furthermore, we considered here only benchmark cases of perception rules for optimal preference intensities. A more general approach would be sensible.

Note also that we restricted the parental decision problems to the socialization side only — and left other behavioral determinants (like general social interactions) unconsidered. Accounting for a richer 'adult world' could yield qualitatively different results. Finally, the role of the children in their socialization process is so far that of passive receivers. Allowing for a pro–active role of children in the adoption process of preferences could also constitute a fruitful extension of the present baseline model (and that of Pichler [16]).

A Proofs

Many parts of the proofs below follow straightforwardly from the general characteristics of parental best reply choices shown in Proposition 1 in Pichler [16] (these characteristics must also hold for the individual best reply choices in a SNE). For ease of reference, we replicate this proposition here. We will be using the notation of section 2, and additionally require the following. For any $a \in A$, we will denote any pair of best reply choices (which are chosen against the representative DPI and subject to the optimally perceived PI, adopted and inter–generational PI) as $\left(\phi_a^d\left(\phi_{A_a}^d, \hat{\phi}_{\tilde{a}}(R_a), \phi_a, i_a\right), \hat{\sigma}_a\left(\phi_{A_a}^d, \hat{\phi}_{\tilde{a}}(R_a), \phi_a, i_a\right)\right)$, which we will abbreviate below as $\left(\phi_a^d(\cdot), \hat{\sigma}_a(\cdot)\right)$. Furthermore, the resulting best reply location of the adult child's adopted PI will be denoted $\phi_{\tilde{a}}\left(\phi_a^d(\cdot), \hat{\sigma}_a(\cdot), \phi_{A_a}^d\right)$.

Proposition A.1 (Characterization of Best Replies)**.** *Let Assumptions 1–3 hold. Then, if*

(a) $\phi_{A_a}^d \neq \hat{\phi}_{\tilde{a}}(R_a)$*, generically*[17] $\operatorname{sign}\left(\phi_a^d(\cdot) - \phi_a\right) = -\operatorname{sign}\left(\phi_{A_a}^d - \hat{\phi}_{\tilde{a}}(R_a)\right)$ *and* $\hat{\sigma}_a(\cdot) > 0$*, while always* $\operatorname{sign}\left(\phi_{\tilde{a}}\left(\phi_a^d(\cdot), \hat{\sigma}_a(\cdot), \phi_{A_a}^d\right) - \hat{\phi}_{\tilde{a}}(R_a)\right) = \operatorname{sign}\left(\phi_{A_a}^d - \hat{\phi}_{\tilde{a}}(R_a)\right)$.

[17] There are two kinds of exceptions to the generic characterization. The first is that if the deviation of the best reply DPI from the adopted PI into the characterized direction is not possible, i.e. if the adopted PI of a parent coincides with (the relevant) one of the boundaries of $\phi^d(X)$, then the best reply DPI will coincide with that boundary (while as still generically $\hat{\sigma}_a(\cdot) > 0$). The second is that in the cases where $\hat{\phi}_{\tilde{a}}(R_a) > \phi_a$ and $\phi_{A_a}^d \in \left(\phi_a, \hat{\phi}_{\tilde{a}}(R_a)\right)$, respectively where $\hat{\phi}_{\tilde{a}}(R_a) < \phi_a$ and $\phi_{A_a}^d \in \left(\hat{\phi}_{\tilde{a}}(R_a), \phi_a\right)$, it can also hold that $\operatorname{sign}\left(\phi_a^d(\cdot) - \phi_a\right) = 0$ and $\hat{\sigma}_a(\cdot) = 0$, hence $\phi_{\tilde{a}}\left(\phi_a, 0, \phi_{A_a}^d\right) = \phi_{A_a}^d$.

(b) $\phi^d_{A_a} = \hat{\phi}_{\tilde{a}}(R_a)$, *it holds that* $\phi^d_a(\cdot) - \phi_a = 0$ *and* $\hat{\sigma}_a(\cdot) = 0$, *hence* $\phi_{\tilde{a}}\left(\phi_a, 0, \hat{\phi}_{\tilde{a}}(R_a)\right) - \hat{\phi}_{\tilde{a}}(R_a) = 0.$

Proof. Confer the proof of Proposition 1 in Pichler [16].

A.1 Proof of Proposition 2

That Proposition 2 holds follows as an immediate consequence of the Lemma below. This shows the range of the phase vectors depicted in Figure 1.

Lemma 1 (Phase Vectors)**.**

(a) If $\phi_{H_t} \geq e_H$ *and* $\phi_{H_t} \geq \phi_{L_t}$ *then* $\dot{\phi}_{H_t} < 0$*; and if* $\phi_{L_t} \leq e_L$ *and* $\phi_{L_t} \leq \phi_{H_t}$ *then* $\dot{\phi}_{L_t} > 0$.

(b) If $\phi_{L_t} \geq \phi_{H_t} \geq e_H$ *then* $\dot{\phi}_{L_t} < 0$*; and if* $e_L \geq \phi_{L_t} \geq \phi_{H_t}$ *then* $\dot{\phi}_{H_t} > 0$.

(c) If $\phi_{L_t} \geq e_L$, $\phi_{L_t} \geq \phi_{H_t}$ *and* $\phi_{H_t} \leq e_H$ *then* $\dot{\phi}_{H_t} > \dot{\phi}_{L_t}$.

Proof. Before we start this proof, note that in the notation of the SNE quantities below, their dependence on P_t is not indicated for brevity. Also, in a slight notational inconsistency, we will denote with $\phi^d_{g_t}\left(\phi^{d^*}_{A_t}\right)$ the best reply choices to the location of the average (SNE–)DPI of any individual member of the cultural groups $G_t \in \{L_t, H_t\}$ (again, the dependence on other parameters is not indicated). Note also that all claims that are made below without comment follow directly from Proposition A.1.

(a) Let $\phi_{H_t} \geq e_H$ and $\phi_{H_t} \geq \phi_{L_t}$. Then $\dot{\phi}_{H_t} < 0$ if $\phi^{d^*}_{A_t} < \phi_{H_t}$. Suppose, by ways of contradiction, that $\phi^{d^*}_{A_t} \geq \phi_{H_t}$. In this case, $\phi^d_{h_t}\left(\phi^{d^*}_{A_t}\right) \leq \phi_{H_t}$. Furthermore, $\phi^d_{l_t}\left(\phi^{d^*}_{A_t}\right) > \phi_{L_t}$ if ϕ_{L_t} is larger than the lower bound of $\phi^d(X)$. If ϕ_{L_t} equals that lower bound, then $\phi_{H_t} > \phi^d_{l_t}\left(\phi^{d^*}_{A_t}\right) = \phi_{L_t}$. This contradicts $\phi^{d^*}_{A_t} \geq \phi_{H_t}$ being supported by best reply choices.

The proof for the 'opposite' case of $\phi_{L_t} \leq e_L$ and $\phi_{L_t} \leq \phi_{H_t}$ is analogous.

(b) Let $\phi_{L_t} \geq \phi_{H_t} \geq e_H$. Then $\dot{\phi}_{L_t} < 0$ if $\phi^{d^*}_{A_t} < \phi_{L_t}$. Suppose, again by ways of contradiction, that $\phi^{d^*}_{A_t} \geq \phi_{L_t}$. In this case, $\phi^d_{l_t}\left(\phi^{d^*}_{A_t}\right) < \phi_{L_t}$ and $\phi^d_{h_t}\left(\phi^{d^*}_{A_t}\right) < \phi_{H_t}$. This yields a contradiction.

The proof for the 'opposite' case of $e_L \geq \phi_{L_t} \geq \phi_{H_t}$ is again analogous.

(c) Let $\phi_{L_t} \geq e_L$, $\phi_{L_t} \geq \phi_{H_t}$ and $\phi_{H_t} \leq e_H$. Then, if $\phi^{d^*}_{A_t} \in (\phi_{H_t}, \phi_{L_t})$ it follows that $\dot{\phi}_{H_t} > 0$ and $\dot{\phi}_{L_t} < 0$. If $\phi^{d^*}_{A_t} \notin (\phi_{H_t}, \phi_{L_t})$, then this can only hold if $\phi^{d^*}_{L_t} < \phi^{d^*}_{A_t} < \phi^{d^*}_{H_t}$. Thus, $\phi_{H_{t+1}} > \phi_{L_{t+1}} \Longrightarrow \dot{\phi}_{H_t} > \dot{\phi}_{L_t}$. □

A.2 Proof of Proposition 3

By Proposition A.1, it suffices to show here that $\phi^{d^*}_{A_t} \in (e_L, e_H)$. Suppose, by ways of contradiction, that $\phi^{d^*}_{A_t} \geq e_H$. But then, $\phi^d_{h_t}\left(\phi^{d^*}_{A_t}\right) \leq \phi_{H_t}$ while $\phi^d_{l_t}\left(\phi^{d^*}_{A_t}\right) < \phi_{L_t}$. This contradicts $\phi^{d^*}_{A_t} \geq \phi_{H_t}$ being supported by best reply choices. The analogous logic yields a contradiction for the case of $\phi^{d^*}_{A_t} \leq \phi_{L_t}$. □

A.3 Proof of Proposition 4

First note that the dynamical system is time–autonomous. For this reason, we will subsequently drop all time–indexes for ease of notation. Further, we will denote the phase vectors corresponding to any point $(\phi_L, \phi_H) \in \phi^d(X)^2$ as $\left(\dot{\phi}_L(\phi_L, \phi_H), \dot{\phi}_L(\phi_L, \phi_H)\right)$ and suppress the dependence on the other (constant) parameters. Also note that by the continuity of the SNE choices, the phase vectors must be continuous in state space.

Consider the point (e_L, e_H). We know (from the proof of Proposition 2) that at this point, $\dot{\phi}_H(e_L, e_H) < 0$. Conversely at the point (e_L, e_L) $\dot{\phi}_H(e_L, e_H) > 0$. It follows that on the 'lower border (LB)' of the generic state space where $\phi_L = e_L$ the signs of $\dot{\phi}_H$ must switch an odd number of times. Similarly, since $\dot{\phi}_L(e_H, e_H) < 0$, it follows that on the 'upper bound (UB)' of the generic state space where $\phi_H = e_H$, the signs of $\dot{\phi}_L$ must switch an odd number of times.

Consider now any of the isoclines that starts at a 'switching point' of $\dot{\phi}_H$ on the LB. Thus, this isocline must feature $\dot{\phi}_H = 0$, and can furthermore not end at the UB, since there $\dot{\phi}_H(\phi_L, e_H) < 0$. Assume now that such an isocline 'starts' and 'ends' at the LB. Since no isoclines where $\dot{\phi}_H = 0$ can cross each other, it follows that between the corresponding start– and endpoints on the LB there must lie (either zero or) an even number of points where $\dot{\phi}_H$ changes signs. It then follows that a $\dot{\phi}_H = 0$–isocline that starts at the LB and ends at the main diagonal (MD) must exist. Furthermore, at least one such isocline must also have 'odd order' on the LB. With this we mean that both counting from above or from below, the number of appearance of the corresponding 'switching point' of $\dot{\phi}_H$ is odd. Finally, on any such isoclines, the sign of $\dot{\phi}_L$ must switch an odd number of times. This follows since on the MD, $\dot{\phi}_H > \dot{\phi}_L$ must hold, thus $\dot{\phi}_L < 0$ at the end–point on the MD of the isocline under scrutiny. Thus, there is an odd number of intersections with isoclines where $\dot{\phi}_L = 0$ (of which an odd number exists).

Since on the LB, $\dot{\phi}_L(e_L, \phi_H) > 0$ it follows from the analogous reasoning that at least one $\dot{\phi}_L = 0$–isocline that starts at the UB and ends at the MD with 'odd order' must exist. Also, the the sign of $\dot{\phi}_H$ must switch an odd number of times thereon (since on its end–point on the MD, $\dot{\phi}_H > 0$ must hold). Thus, there is an odd number of intersections with isoclines where

$\dot{\phi}_H = 0$ (of which an odd number exists).

From these properties, it is now immediate that at least one pair of the above described 'odd order' isoclines must intersect — where the intersection can only take place at 'odd order' (on both isoclines). Consider now this intersection, which obviously constitutes a steady state. Denote this (ϕ_L^*, ϕ_H^*). Now, since the intersection can only take place at 'odd order', it follows that there is a nonempty neighborhood around the steady state where (a) for all points on the respective $\dot{\phi}_H = 0$–isocline such that $\phi_L <=> \phi_L^*$, it holds that $\dot{\phi}_L (\phi_L, \phi_H) >=< 0$, and (b) for all for all points on the respective $\dot{\phi}_L = 0$–isocline such that $\phi_H <=> \phi_H^*$, it holds that $\dot{\phi}_H (\phi_L, \phi_H) >=< 0$. By the continuity of the phase vectors it then finally follows that the steady state (ϕ_L^*, ϕ_H^*) is locally asymptotically stable. □

A.4 Proof of Proposition 5

Consider any $\phi_{H_t} > \phi_{L_t}$. Further, denote with the pair $\left(\phi_{L_t}^d, \phi_{H_t}^d\right)$ the (average) DPIs of the members of the cultural groups (these coincide with the parents' optimal PIs). It then follows that $\phi_{A_t}^d = (1 - q_H)\phi_{L_t}^d + q_H \phi_{H_t}^d$. Assume now that $\phi_{H_t}^d \geq \phi_{H_t} > \phi_{L_t} \geq \phi_{L_t}^d$, in which case $\phi_{A_t}^d \in \left(\phi_{L_t}^d, \phi_{H_t}^d\right)$.

From Proposition A.1, we know that for any of the individual best replies of the members of both groups to such a constellation, it must hold that $\phi_{h_t}^d \left(\phi_{L_t}^d, \phi_{H_t}^d\right) \geq \phi_{H_t}$ and $\phi_{l_t}^d \left(\phi_{L_t}^d, \phi_{H_t}^d\right) \leq \phi_{L_t}$. The sets of individual best reply DPIs are furthermore non–empty, compact and UHC by Berge's Theorem of the Maximum. By the concavity of the target functions of the individual optimization problems, these sets are furthermore convex.

We thus have a non–empty, convex– and compact–valued and UHC correspondence from a non–empty, convex and compact set into itself. By Kakutani's Fixed Point Theorem, there must thus exist a fixed point of the form

$$\left(\phi_{l_t}^d \left(\phi_{L_t}^d, \phi_{H_t}^d\right), \phi_{h_t}^d \left(\phi_{L_t}^d, \phi_{H_t}^d\right)\right) = \left(\phi_{L_t}^d, \phi_{H_t}^d\right).$$

From Proposition A.1, we know additionally that, under the conditions above, whenever ϕ_{H_t}/ϕ_{L_t} is strictly smaller/larger than the upper/lower bound of the set of possible DPIs, then (given the described constellation above) any of the individual best replies must feature $\phi_{h_t}^d \left(\phi_{L_t}^d, \phi_{H_t}^d\right) > \phi_{H_t}$ and $\phi_{l_t}^d \left(\phi_{L_t}^d, \phi_{H_t}^d\right) < \phi_{L_t}$, and $\hat{\sigma}_{g_t} \left(\phi_{L_t}^d, \phi_{H_t}^d\right) > 0$, $\forall g_t \in \{l_t, h_t\}$. This finalizes the proof of part (a) of Proposition 5.

To see part (b) simply note that if $\phi_{H_t} = \phi_{L_t} = \phi$, then it must hold that $\left(\phi_{l_t}^d (\phi, \phi), \phi_{h_t}^d (\phi, \phi)\right) = (\phi, \phi)$ and $(\hat{\sigma}_{l_t} (\phi, \phi), \hat{\sigma}_{h_t} (\phi, \phi)) = (0, 0)$. □

A.5 Proof of Proposition 6

Again due to the time–autonomy of the dynamical system, we will drop the time–indexes in the subsequent proof.

(a) Consider any arbitrary point on the main diagonal of the state space, (ϕ, ϕ). Then, by the continuity of the SNE choices, there must exist an open neighborhood around this point, $N_\epsilon(\phi, \phi)$, with $\phi_L \neq \phi_H$, $\forall (\phi_L, \phi_H) \in N_\epsilon(\phi, \phi)$, and for which it holds that $0 < \hat{\sigma}^*_G (\phi_L, \phi_H) < \epsilon < 1$, $\forall G = \{L, H\}$.

Assume now, by ways of contradiction, that the distance between the PIs of the two cultural groups would be increasing (i.e. not strictly decreasing) at any point in this open neighborhood. Using the previous inequality, it follows from equation (5) that $\forall (\phi_L, \phi_H) \in N_\epsilon(\phi, \phi)$,

$$\frac{\left|\phi^{d^*}_L (\phi_L, \phi_H) - \phi^{d^*}_H (\phi_L, \phi_H)\right|}{|\phi_L - \phi_H|} > \frac{1}{\epsilon} > 1$$

would have to hold. Now, remember from Proposition 5 that $\phi^{d^*}_G (\phi, \phi) = \phi$, $\forall G = \{L, H\}$. Thus, for any sequence in $N_\epsilon(\phi, \phi)$ that converges to the point (ϕ, ϕ), it would have to hold by the continuity of the DPI choices that the limit of the left hand side of the above inequality equals one. But this yields a contradiction. □

(b) This proof will be based on additional normalizations of the SNE selection function. To discuss these, we will focus our attention to the upper triangle of the state space (i.e. where $\phi_H \geq \phi_L$).

Consider a point where the adopted PI of the members of group L coincides with the lower bound of $\phi^d(X) = \left[\underline{\phi}, \overline{\phi}\right]$. Denote this $\left(\underline{\phi}, \phi^1_H\right)$, with $\phi^1_H > \underline{\phi}$. We know that at any such point $\phi^{d^*}_L \left(\underline{\phi}, \phi^1_H\right) = \underline{\phi}$, and the DPI constraint is binding. Then, there is always a SNE selection function for which it holds that there is a non–empty and right–open interval $\left[\left(\underline{\phi}, \phi^1_H\right), \left(\phi^1_L, \phi^1_H\right)\right)$, where $\phi^1_L > \underline{\phi}$, for which it holds that (a) the DPI constraint for L stays binding; as well as that (b) for all points $\left(\phi_L, \phi^1_H\right)$ in this interval, $\phi^{d^*}_H \left(\phi_L, \phi^1_H\right) = \phi^{d^*}_H \left(\underline{\phi}, \phi^1_H\right)$ and $\hat{\sigma}^*_H \left(\phi_L, \phi^1_H\right) = \hat{\sigma}^*_H \left(\underline{\phi}, \phi^1_H\right)$.

Extending this sort of normalization to any point where $\phi_L = \underline{\phi}$ (and which is not located on the main diagonal), we obtain a continuum of right–open intervals on any of which it holds that $\dot{\phi}_H$ is constant.

Analogously, consider a point where $\phi_H = \overline{\phi}$, and denote this point $\left(\phi^2_L, \overline{\phi}\right)$, with $\phi^2_L < \overline{\phi}$. We know that at this point $\phi^{d^*}_H \left(\phi^2_L, \overline{\phi}\right) = \overline{\phi}$, and the DPI constraint is binding. Then, there is always a SNE selection function for which it holds that there is a non–empty and left–open interval $\left(\left(\phi^2_L, \phi^2_H\right), \left(\phi^2_L, \overline{\phi}\right)\right]$, where $\phi^2_H < \overline{\phi}$, for which it holds that (a) the DPI constraint for H stays binding; as well as that (b) for all points $\left(\phi^2_L, \phi_H\right)$ in this interval, $\phi^{d^*}_L \left(\phi^2_L, \phi_H\right) = \phi^{d^*}_L \left(\phi^2_L, \overline{\phi}\right)$ and $\hat{\sigma}^*_L \left(\phi^2_L, \phi_H\right) = \hat{\sigma}^*_L \left(\phi^2_L, \overline{\phi}\right)$.

Extending this sort of normalization to any point where $\phi_H = \overline{\phi}$ (and which is not located on the main diagonal), we obtain a continuum of left–open intervals on any of which it holds that $\dot{\phi}_L$ is constant.

Consider now any pair of points that consists of a right–boundary point of the first type of intervals and a left–boundary point of the second type of intervals, which satisfies the following conditions: (a) At the first of these points, the DPI constraint for group H is not binding, and at the second of these points, the DPI constraint for group L is not binding. (b) These two points are connected through a $45\,^{\circ}$–line–segment (in state space).

We will now impose a normalization on the SNE selection on these sorts of $45\,^{\circ}$–line–segments. To introduce this, the following definition will be useful.

Definition A.1 (State–corrected SNE Choices). *$\forall\,(\phi_L, \phi_H) \in \operatorname{con}\phi^d(X)^2$, denote the tuple*

$$\left\{\phi_G^{d^*}\left(\phi_L, \phi_H\right) - \phi_G, \hat{\sigma}_G^*\left(\phi_L, \phi_H\right)\right\}_{G\in\{L,H\}}$$

as state–corrected SNE choices*.*

We will now indeed require the SNE selection function to select identical state–corrected SNE choices for every point on any of the above constructed $45\,^{\circ}$–line–segments. Thus for all points on such $45\,^{\circ}$–line–segments, both $\dot{\phi}_L$ and $\dot{\phi}_H$ are constant (i.e. the $45\,^{\circ}$–line–segments are isoclines).

We can now give the following summarizing characterization of the phase vectors in the upper triangle of the state space. First, the main diagonal consists of a continuum of steady states (Proposition 5 (b)). This is neighbored by a continuum of line–segments consisting of a connection of (a) a horizontal line in state space where $\dot{\phi}_H$ is constant, with (b) a $45\,^{\circ}$–isocline, with (c) a vertical line where $\dot{\phi}_L$ is constant. The 'upper' border of this neighboring field of the main diagonal is as follows. There is a point (ϕ^m, ϕ^m) where $\left(\phi_L^{d^*}\left(\phi^m, \phi^m\right), \left(\phi_H^{d^*}\left(\phi^m, \phi^m\right)\right)\right) = \left(\underline{\phi}, \overline{\phi}\right)$, but where both DPI constraints are unbinding. This single point constitutes its own $45\,^{\circ}$–line–segment. Thus, the vertical and horizontal lines connected to this point do constitute the borders of a rectangle in which it holds that the lower DPI constraint is binding for group L and the upper DPI constraint is binding for group H. This rectangle is thus made of a continuum of horizontal lines with constant $\dot{\phi}_H$ and a continuum of vertical lines with constant $\dot{\phi}_L$.

From these properties, it follows straightforwardly that no cycles can exist in the upper triangle of the state space. Thus, all sequences of PIs must converge to a steady state therein. Extending these properties (respectively the normalizations on the SNE selection function) to the lower triangle of the state space in an analogous way, we obtain the convergence property for the whole state space. □

B Extensions

B.1 Evolution under Imperfect Empathy

The central property of the evolution under global 'imperfect empathy' (respectively the first type of perception rule) is that if the oblique socialization is unbiased, then the PIs of (almost) all adults converge to the same point (confer Proposition 4 in Pichler [16]). Since in the present paper, we have assumed unbiased oblique socialization, this result must thus hold.

It rests to show the characterization of the SNE choices for the two cultural groups case and under global imperfect empathy.

Proposition B.1 (Characterization of SNE choices)**.** *Let Assumptions 1–5 hold. Then, the following SNE properties are satisfied,* $\forall\{i_L, i_H, q_H\} \in (\mathbb{R}_{++} \setminus \{\infty\})^2 \times (0,1)$.

1. *Case* $\phi_{H_t} <> \phi_{L_t}$
   (a) $\phi^{d^*}_{H_t}(P_t) <> \phi_{H_t} <> \phi_{H_{t+1}} <> \phi_{L_{t+1}} <> \phi_{L_t} <> \phi^{d^*}_{L_t}(P_t)$[18]*, and*
   (b) $\hat{\sigma}^*_{G_t}(P_t) \in (0,1)$, $\forall G_t \in \{L_t, H_t\}$.
2. *Case* $\phi_{H_t} = \phi_{L_t}$
   (a) $\phi^{d^*}_{G_t}(P_t) = \phi_{G_t} = \phi_{G_{t+1}}$, $\forall G_t \in \{L_t, H_t\}$*, and*
   (b) $\hat{\sigma}^*_{G_t}(P_t) = 0$, $\forall G_t \in \{L_t, H_t\}$.

Proof. Let $\phi_{H_t} > \phi_{L_t}$. By Proposition A.1, it suffices to show that $\phi^{d^*}_{A_t}(P_t) \in (\phi_{L_t}, \phi_{H_t})$. Assume, by ways of contradiction, that $\phi^{d^*}_{A_t}(P_t) \geq \phi_{H_t}$. However, in this case $\phi_{h_t}\left(\phi^{d^*}_{A_t}(P_t)\right) \leq \phi_{H_t}$ while even $\phi_{l_t}\left(\phi^{d^*}_{A_t}(P_t)\right) < \phi_{H_t}$. This yields a contradiction. Analogously, we obtain a contradiction for $\phi^{d^*}_{A_t}(P_t) \leq \phi_{L_t}$. Also, the proof for the case $\phi_{H_t} > \phi_{L_t}$ is analogous.

Let $\phi_{H_t} = \phi_{L_t}$, and assume that $\phi^{d^*}_{A_t}(P_t) <> \phi_{H_t} = \phi_{L_t}$. But then, $\phi_{g_t}\left(\phi^{d^*}_{A_t}(P_t)\right) >< \phi^{d^*}_{A_t}(P_t)$, $\forall g_t \in \{l_t, h_t\}$, which yields a contradiction again. □

By the results of Proposition B.1, the cultural groups strictly assimilate inter–generationally. Thus, the distance between the PIs of the two groups is a contraction mapping and the PIs of the groups converge to the same point (confirming the result of Pichler [16]). This result can be interpreted to correspond to the 'melting pot' theory of assimilation of cultural groups (see e.g. Han [13]).

B.2 Comparative Statics

For pursuing the comparative statics subject to the types of perception rules that we considered in the present paper, the following additional assumptions are necessary.

Assumption B.1 (Curvature)**.**

[18] Again, the outer inequalities would be strict if the respective adopted PI would coincide with the relevant boundary of $\operatorname{con}\phi^d(X)$. But this can only be the case in the initial period, given the results of the present Proposition.

(a) u^b, $v^d(\cdot\,|h)$ *and* c *are* $\mathcal{C}^2$*, and*

(b) $\operatorname{sign}(d-k)\,\frac{\partial^2 v^d(k|h)}{\partial k\,\partial h} > 0,\ \forall\,(k,h) \in \operatorname{con}\phi^d(X) \times \mathbb{R}_{++}$.

Assumption B.1 (b) means that the marginal cost of a deviation of the adopted PI of the adult child from the optimal PI is strictly increasing in the inter–generational PI. Notably, this is only necessary for the results of the third and fourth column of the comparative statics matrix below to hold.

The results of the following Proposition are valid for all specifications of the perception rules that we consider in the present paper.

Proposition B.2 (Comparative Statics)**.** *Let Assumptions 1–B.1 be satisfied. Assume that* $\phi_{H_t} > \phi_{L_t}$*,*[19] *and let us distinguish the cultural groups as* $G_t = L_t/H_t$*. Then, if the optimization problems of the parents of both groups are strictly concave at at the SNE, and if the two decision variables are 'not too strong substitutes', the following comparative statics results hold*

$$\operatorname{sign}\begin{pmatrix} \frac{\partial\,\phi^{d^*}_{G_t}(P_t)}{\partial\,\phi_{G_t}} & \frac{\partial\,\phi^{d^*}_{G_t}(P_t)}{\partial\,\phi_{-G_t}} & \frac{\partial\,\phi^{d^*}_{G_t}(P_t)}{\partial\,i_G} & \frac{\partial\,\phi^{d^*}_{G_t}(P_t)}{\partial\,i_{-G}} \\ \frac{\partial\,\sigma^*_{G_t}(P_t)}{\partial\,\phi_{G_t}} & \frac{\partial\,\sigma^*_{G_t}(P_t)}{\partial\,\phi_{-G_t}} & \frac{\partial\,\sigma^*_{G_t}(P_t)}{\partial\,i_G} & \frac{\partial\,\sigma^*_{G_t}(P_t)}{\partial\,i_{-G}} \end{pmatrix} = \begin{pmatrix} +1/+1 & -1/-1 & -1/+1 & -1/+1 \\ -1/+1 & +1/-1 & +1/+1 & +1/+1 \end{pmatrix}$$

Proof. Below, we will denote with $\mathcal{L}^*_G$ the identical Lagrangeans of the optimization problems of all parents of any group $G_t \in \{L_t, H_t\}$. For saving space, we will mostly drop the time–indexes below. By the Implicit Function Theorem[20], we obtain the following comparative statics effects.

$$\operatorname{sign}\left(\frac{\partial\,\phi^{d^*}_G}{\partial\,\phi_G}\right) = \operatorname{sign}\left(\left(\left(\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat\sigma_{\text{-G}}}\right)^2 - \frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,{\phi^{\text{d}}_{\text{-G}}}^2}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,{\hat\sigma_{\text{-G}}}^2}\right)\left(\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\phi_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,{\hat\sigma_{\text{G}}}^2} - \frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat\sigma_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat\sigma_{\text{G}}\,\partial\,\phi_{\text{G}}}\right)\right)$$

$$\begin{aligned}\operatorname{sign}\left(\frac{\partial\,\sigma^*_G}{\partial\,\phi_G}\right) = {} & \\ \operatorname{sign}\Bigg(&\left(\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,{\phi^{\text{d}}_{\text{-G}}}^2}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,{\hat\sigma_{\text{-G}}}^2} - \left(\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat\sigma_{\text{-G}}}\right)^2\right)\left(\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\phi_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat\sigma_{\text{G}}} - \frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,{\phi^{\text{d}}_{\text{G}}}^2}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat\sigma_{\text{G}}\,\partial\,\phi_{\text{G}}}\right) + \\ &\left(\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,{\phi^{\text{d}}_{\text{G}}}^2}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,{\hat\sigma_{\text{G}}}^2} - \left(\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat\sigma_{\text{G}}}\right)^2\right)\left(\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat\sigma_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat\sigma_{\text{-G}}\,\partial\,\phi^{\text{d}}_{\text{G}}} - \frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\phi^{\text{d}}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,{\hat\sigma_{\text{-G}}}^2}\right)\Bigg)\end{aligned} \tag{B.1}$$

[19]This normalization is without loss of generality. The results for the opposite case would be identical, with the comparative statics signs reversed.

[20]Since the parental optimization problems are strictly concave at any SNE under consideration, the determinant of the Hessian matrix is strictly positive and all conditions for the Implicit Function Theorem are fulfilled.

$$\text{sign}\left(\frac{\partial\,\phi_G^{d*}}{\partial\,\phi_{-G}}\right) =$$
$$\text{sign}\left(\left(\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\phi_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}{}^2} - \frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}\,\partial\,\phi_{\text{-G}}}\right)\left(\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}{}^2} - \frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat{\sigma}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\right)\right)$$

$$\text{sign}\left(\frac{\partial\,\sigma_G^{*}}{\partial\,\phi_{-G}}\right) =$$
$$\text{sign}\left(\left(\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}\,\partial\,\phi_{\text{-G}}} - \frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\phi_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}{}^2}\right)\left(\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat{\sigma}_{\text{G}}} - \frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}{}^2}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\right)\right)$$

$$\text{sign}\left(\frac{\partial\,\phi_G^{d*}}{\partial\,\phi_G^{s}}\right) =$$
$$\text{sign}\left(\left(\left(\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}\right)^2 - \frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}{}^2}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}{}^2}\right)\left(\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\text{i}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}{}^2} - \frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat{\sigma}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\text{i}_{\text{G}}}\right)\right)$$

$$\text{sign}\left(\frac{\partial\,\sigma_G^{*}}{\partial\,\phi_G^{s}}\right) = \tag{B.2}$$
$$\text{sign}\Bigg(\left(\left(\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}\right)^2 - \frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}{}^2}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}{}^2}\right)\left(\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}{}^2}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\text{i}_{\text{G}}} - \frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat{\sigma}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\text{i}_{\text{G}}}\right) +$$
$$\left(\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\phi^{\text{d}}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}{}^2} - \frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}\,\partial\,\phi^{\text{d}}_{\text{G}}}\right)\left(\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\text{i}_{\text{G}}} - \frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\text{i}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\right)\Bigg)$$

$$\text{sign}\left(\frac{\partial\,\phi_G^{d*}}{\partial\,\phi^s_{-G}}\right) =$$
$$\text{sign}\left(\left(\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\text{i}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}{}^2} - \frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}\,\partial\,\text{i}_{\text{-G}}}\right)\left(\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}{}^2} - \frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat{\sigma}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\right)\right)$$

$$\text{sign}\left(\frac{\partial\,\sigma_G^{*}}{\partial\,\phi^s_{-G}}\right) =$$
$$\text{sign}\left(\left(\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}\,\partial\,\text{i}_{\text{-G}}} - \frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\text{i}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}{}^2}\right)\left(\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat{\sigma}_{\text{G}}} - \frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}{}^2}\,\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\right)\right)$$

To analyze the comparative statics effects in the sign–equations above, first note that all bracket–terms that do include neither one of the second partial derivatives $\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat{\sigma}_{\text{G}}}$ or $\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}$ are unambiguous in sign. Since the sign of $\frac{\partial^2\,\mathcal{L}^*_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat{\sigma}_{\text{G}}}$ and $\frac{\partial^2\,\mathcal{L}^*_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}$ are ambiguous, this also holds for all bracket–terms above where these expressions are included. For $G = L/H$, consider the vector

$$b_G := \begin{pmatrix} \frac{\frac{\partial^2 \mathcal{L}^*_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \, \partial \phi_{\mathrm{G}}}}{\frac{\partial^2 \mathcal{L}^*_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \, \partial \phi_{\mathrm{G}}}} & \frac{\frac{\partial^2 \mathcal{L}^*_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \, \partial \mathrm{i}_{\mathrm{G}}}}{\frac{\partial^2 \mathcal{L}^*_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \, \partial \mathrm{i}_{\mathrm{G}}}} & \frac{\frac{\partial^2 \mathcal{L}^*_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \, \partial \phi^{\mathrm{d}}_{\mathrm{-G}}}}{\frac{\partial^2 \mathcal{L}^*_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \, \partial \phi^{\mathrm{d}}_{\mathrm{-G}}}} \end{pmatrix}$$

and note that for cultural group L/H, all entries of this matrix are strictly negative/positive. Next, let $\forall G \in \{L, H\}$

$$\frac{\partial^2 \mathcal{L}^*_G}{\partial \phi^d_G \, \partial \hat{\sigma}_G} < / > (\max/\min \, b_G) \frac{\partial^2 \mathcal{L}^*_G}{\partial {\phi^d_G}^2} \text{ and } \frac{\partial^2 \mathcal{L}^*_G}{\partial \phi^d_G \, \partial \hat{\sigma}_g} < / > \left(\max/\min \, \frac{1}{b_G}\right) \frac{\partial^2 \mathcal{L}^*_G}{\partial \hat{\sigma}^2_G} \tag{B.3}$$

This basically requires that the two socialization instruments must not be too strong substitutes in the parental optimization problem (3) (or that the optimization problem must be sufficiently concave in both socialization instruments compared to the cross–concavity). It is then straightforward to show that under the conditions (B.3), the comparative statics effects of Proposition B.2 hold. □

References

[1] Bisin, A. and Topa, G. (2003). Empirical models of cultural transmission. *J. Europ. Econ. Ass.*, 1(2-3):363–375.

[2] Bisin, A., Topa, G., and Verdier, T. (2009). Cultural transmission, socialization and the population dynamics of multiple-trait distributions. *Int. J. Econ. Theory*, 5(1):139–154.

[3] Bisin, A. and Verdier, T. (1998). On the cultural transmission of preferences for social status. *J. Pub. Econ.*, 70(1):75–97.

[4] Bisin, A. and Verdier, T. (2000). A model of cultural transmission, voting and political ideology. *Europ. J. Polit. Economy*, 16(1):5–29.

[5] Bisin, A. and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *J. Econ. Theory*, 97(2):298–319.

[6] Bisin, A. and Verdier, T. (2010). The economics of cultural transmission and socialization. In *Handbook of Social Economics (forthc.)*. Elsevier Science.

[7] Boyd, R. and Richerson, P. J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press.

[8] Cavalli-Sforza, L. L. and Feldman, M. (1973). Cultural versus biological inheritance: phenotypic transmission from parents to children. *Am. J. Hum. Genet.*, 25:618–637.

[9] Cavalli-Sforza, L. L. and Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach.* Princeton University Press, Princeton.

[10] Cox, D. and Stark, O. (1996). Intergenerational transfers and the demonstration effect. Boston College Working Papers in Economics 329., Boston College Department of Economics.

[11] Festinger, L. (1957). *A theory of cognitive dissonance.* Evanston, IL: Row, Peterson.

[12] Grusec, J. E. (2002). *Handbook of Parenting: Practical Issues in Parenting*, chapter Parental Socialization and Children's Acquisition of Values, pages 153–167. Lawrence Erlbaum Associates.

[13] Han, P. (2006). *Theorien zur internationalen Migration.* Lucius & Lucius.

[14] Otto, S. P., Christiansen, F. B., and Feldman, M. W. (1994). Genetic and cultural inheritance of continuous traits. Morrison institute for population and resource studies working paper no. 0064, Stanford University.

[15] Panebianco, F. (2010). "driving while black": A theory for interethnic integration and evolution of prejudice. Eurodiv Paper 70.2010.

[16] Pichler, M. M. (2010). The economics of cultural formation of preferences. Working Papers 431, Bielefeld University, Institute of Mathematical Economics.

[17] Rath, K. P. (1992). A direct proof of the existence of pure strategy equilibria in games with a continuum of players. *Econ. Theory*, 2(3):427–33.

[18] Schmeidler, D. (1973). Equilibrium points in nonatomic games. *J. Stat. Phys.*, 7(4).

[19] Stark, O. (1995). *Altruism and Beyond.* Cambridge University Press.