Culture Formation and Endogenous Cultural Distance

Michael M. Pichler*
Institute of Mathematical Economics, Bielefeld University
Postfach 100131, D-33501 Bielefeld, Germany
E–mail: mpichler@uni–bielefeld.de
Phone: +49 521 106 2573, Fax: +49 521 106 2997

Second Draft
July 2009

Abstract

Based on the 'economics of cultural transmission' literature, this paper introduces a generalized representation of the formation of continuous cultural values (traits) in the socialization process. We define the culture of a person as a set of cultural values and introduce a cultural formation of preferences framework that models the culture of a child as the collective outcome of all socialization influences that it experiences. Thereby, a socialization influence is constituted by the chosen cultural behavior (demonstrated culture) of a person and the (relative) success that it has in the socialization process of the child. By restricting the set of agents from which a child socially learns in the socialization period to the adult generation only, we can endogenize the cultural formation of preferences process as resulting out of the optimal parental choices of their demonstrated culture (which constitutes a strategic interaction choice) and their (relative) success in the socialization process of their child. We apply this framework to an OLG environment, where the society is populated by two distinct cultural groups and analyze the endogenous evolution of the distances between the cultures of the groups.

Keywords: Cultural transmission; Socialization; Preference Evolution
JEL–Classification numbers: C72, D10, J13, Z13

---

*This research project has been generously supported by the Deutsche Forschungsgesellschaft (DFG). The author thanks Herbert Dawid and Walter Trockel for invaluable advice. I am further grateful for useful comments from Fabrizio Panebianco and Stefan Zeugner. Special thanks are due to the members of the International Research Training Group 'Economic Behavior and Interaction Models' (EBIM).

1 Introduction

1.1 Motivation

In recent years, the question of assimilation and integration of immigrants with different cultural backgrounds into hosting societies has attained increasing attention, both in media and on the political agenda. In Europe, the debate has centered around an alleged lack of assimilation and integration of Muslim immigrants, compared to other immigrant groups (see e.g. Bisin et al., 2008, who also confirm this allegation empirically, using a U.K. data–set and finding significantly less and slower integration of Muslims compared to non–Muslims). Compared to the European nation states, the debate in the U.S.A. has been less pronounced given its historical pre–disposition as an immigrant country. Still influential is the 'melting pot' theory, which assumes that new immigrants will integrate into the society, thereby contributing to the creation of a new cultural identity (Han, 2006, p. 32). While this theory is based on, and successful to describe, the experience of 'old immigration' from western and northern European immigrants (which are culturally close to anglo-saxian Americans), it fails to explain the tendency towards cultural pluralism that the U.S. society has experienced thereafter (Gordon, 1964, pp. 115–119, 132–136).

The aim of the present paper is to present a theoretical framework that contributes to the understanding of the process of *cultural assimilation* and *behavioral assimilation*. We define cultural assimilation as the inter–temporal decrease in the distance between the cultures of two groups, underlying the concept of *culture* as a set of *cultural values*. Similarly, we interpret behavioral assimilation as the inter–temporal decrease in the distance of the (homogeneous) choices of *demonstrated cultures* of the members of two cultural groups[1]. Thereby, a demonstrated culture is defined as the set of cultural values that a person demonstrates through its socio–economic actions.

The analysis is embedded in a *cultural formation of preferences* framework, in which the culture of a child is formed as the collective outcome of all socialization influences that it experiences. Thereby, a socialization influence is constituted by the demonstrated culture of a person and the (relative) success that it has in the socialization process of the child. By restricting the set of agents, from which a child socially learns in the socialization period to the adult generation only, we can endogenize the cultural formation of preferences process as resulting out of the optimal parental choices of their demonstrated culture (which can also be observed by unrelated children) and their (relative) success in the socialization process of

[1]Cultural assimilation and behavioral assimilation in this context are hence mutual concepts in that they can originate from just one, or both of the cultural groups under scrutiny.

their child.

1.2 Existing Literature

The present analysis stands in a natural context to the existing literature on the *economics of cultural transmission*. This literature is based on the work of Cavalli-Sforza and Feldman (1973, 1981) and Boyd and Richerson (1985) in evolutionary anthropology. Bisin and Verdier (2000, 2001) presented a general framework to study the population dynamics of the distribution of cultural traits under an endogenous intergenerational cultural transmission mechanism. In this framework, which is now standard in the literature, the endogeneity stems from the purposeful parental choice of socialization intensity, which effectively determines the probability that the child will directly adopt the culture 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 culture — which causes subjective utility losses for the parents. The properties of the model framework have been used in the analysis of several different issues, such as e.g. preferences for social status (Bisin and Verdier, 1998), voting and political ideology (Bisin and Verdier, 2000), corruption (Hauk and Sáez-Martí, 2002), hold up problems (Olcina and Penarrubia, 2004).[2] For a more complete overview on the foundations and theoretical and empirical contributions of the literature see Bisin and Verdier (2008).

Related to this strand of literature is the earlier contribution of Stark (1995) who argues that parents might choose altruistic behavior in front of their children even though they are themselves not altruistic. This comes in a (purely egoistic) attempt to instrument the 'demonstration (or preference shaping) effect' of altruistic behavior in order to increase the probability of the child's adoption of this value — and to profit from the child's eventual future care for the parents.

While as the above representations of the cultural transmission or preference adoption process as a probabilistic adoption rule have their merits in the context of discrete traits or preferences, it is less suitable to model the formation of *continuous* traits or preferences. Bisin and Topa (2003) highlighted this issue by reinterpreting the socialization influence of the family as determining its weight in the process of the formation of the traits relative to the weight of the general social environment[3]. The parental motivation

[2] The models mentioned restrict the socialization environment of a child to the adult generation only. See Sáez-Martí and Sjögren (2007) for an analysis on the role that peers play in the socialization process of children.

[3] For an earlier treatment on the issue of inheritance and formation of continuous traits in an empirical context, see Otto et al. (1994). For a treatment of the dynamic evolution of continuous population attributes in a pure social interactions framework see Brueckner and Smirnov (2007, 2008).

to invest in the socialization weight then stems from the aim to avoid subjective utility losses that are increasing in the distance of the child's trait to the parents own trait. Panebianco (2009) uses this framework to study the dynamic evolution of (continuous) inter–ethnic attitudes and the rise of ethnic–based social hierarchies. The focus is on the study of the effect of different oblique socialization schemes on the integration process of ethnic groups. He considers two ethnic groups as fully integrated if they share the same attitudes toward any other ethnic group.

Also recently, Bisin et al. (2006) introduced parental identity–choice[4] into the economics of cultural transmission framework to study the (behavioral) assimilation process of cultural minorities. Parents of the minority group can choose among a continuum of 'life–styles', all of which represent different degrees of assimilation to the social norm of the majority. A lack of assimilation causes social costs that positively depend on the population share of the majority group and negatively on the degree of ethnic identity (which is a decision variable for the parents[5]). The parental problem of choosing the optimal level of assimilation and ethnic identity is nevertheless separated from the socialization problem (i.e. both choices have no social learning impact on the child), which is organized along the standard lines of the economics of cultural transmission literature. The dynamic evolution of the degree of assimilation of the cultural minority then derives from the dynamics of the population distribution of cultures and the evolution of the degree of ethnic identity, which determines the parentally perceived utility cost of the anticipated life–style of the child.

1.3 Paper Contributions

The core of the present paper is the introduction of an *(endogenous) cultural formation of preferences* framework that provides for a general treatment of the formation of continuous cultural values (or preferences or traits) in the socialization process of a child.

Socialization is the process by which individuals learn their culture. This includes the values, norms and attitudes, as well as the beliefs, religious faith, knowledge, customs, language, skills and practices of the society that the individual is embedded in. In the present paper, we focus on the socialization of children with respect to the first three categories mentioned. The reason for this is that they have an inherent continuous structure and can be represented numerically (the importance of a value, the strength of a norm, how positive/negative a certain attitude is) — and can furthermore

[4]The concept of identity formation is based on the work of Akerlof and Kranton (2000).

[5]Bisin et al. (2006) motivate the construction of ethnic identity as a psychological mechanism that is rationally employed to reduce the psychological costs associated with behaving in a non–assimilated way relative to the social norm of the majority cultural group.

be instrumented to re–interpret inherently discrete cultural elements in a continuous way[6]. We will in the present paper collectively refer to the three concepts values, norms and attitudes as *cultural values*[7] and define the full set of cultural values that are subject to the socialization process as the *culture* (that an individual adopts).

Sociological theory emphasizes the fact that the major force of the socialization process is the social learning (of children) from role–models that are demonstrated in the social environment (see e.g. the social learning theories of Bandura and Walters (1963), Bandura (1977), Rotter (1954), Festinger (1957), Skinner (1938, 1953) and others). Within the context of the present paper, the role–model that a person in the social environment of a child represents corresponds to the demonstration of one specific set of cultural values, i.e. one specific culture[8]. We assume that this is realized through the choice of a set of socio–economic actions out of a feasible set of socio–economic actions[9]. For obvious reasons, we will from now on refer to this concept of a role–model as *demonstrated culture*, i.e. the set of *demonstrated cultural values.*

As far as the process of *culture formation* of a child is concerned, we assume that the cultural values that it will adopt through the socialization process realize as the collective outcome of all different demonstrated cultures it socially learns from (as in Bisin and Topa, 2003). Thereby, we restrict the socialization environment of the children to the adult society only. The weight that any of these adults has in the culture formation process depends on the relative success that it had in the socialization of the child

[6]More precisely, values define the valuations that a person assigns to different aspects of life, like the importance of family, friendship, career, or social status. Notably, this concept also includes the valuations of (respectively the importance of) discrete cultural elements like a specific religious faith, the acquisition of certain skills, the conduct of customs and certain practices or the own language. Norms reflect the values of a society or certain groups and define the rules of behavior therein. These rules are transmitted through a system of social rewards (for conformity) and punishment (for non–conformity). Attitudes include the views on socio–economic or political issues, like how much a person is in favor of equal gender treatment, integration of minority groups or an egalitarian society, as well as 'direct attitudes' like the prejudices against members of other cultural groups.

[7]The concept of cultural values is thus closely related to what Sen (1977) terms 'commitments' and Harsanyi (1982) 'moral preferences'.

[8]Stark (1995) discusses a model where parents instrument the 'demonstration effect' of altruistic behavior to increase the probability that children will adopt altruism as a value.

[9]Typical elements of such a set that would correspond to the demonstration of values would include the extensity (time and frequency) and intensity (effort and devotion) of the pursuit of different aspects of life, different ways of self–expression or to achieve and demonstrate status. Elements that would correspond to the demonstration of (the strength) of norms would primarily concern the intensity of the social rewards and punishments for conformity and non–conformity of behavior, while elements that translate into the demonstration of attitudes would be the possible behaviors in different social situations and interactions. Notably, the verbal expression of opinion and thought is a means to demonstrate all categories of cultural values.

compared to the aggregated socialization successes of all other socializing persons.

Given that the culture that a person adopts after the socialization process has been formed by the social learning from role–models, viz. demonstrated cultures, we assume that it induces preferences over choices of demonstrated cultures in the adult life–period of that person. These preferences have the basic property that parents prefer demonstrated cultures that feature individual demonstrated cultural values that are as close as possible to the adopted cultural values (this assumption is in line with the *cognitive dissonance* theory of Festinger, 1957). The interdependence between demonstrated cultures, cultures and preferences induced by cultures constitutes the motivation to label the representation of the socialization process that this paper proposes as *cultural formation of preferences.*

As part of their culture, adults do also adopt *socialization values.* These establish relevance of the future demonstrated culture of their children for their own lives, and can both be of altruistic or egoistic nature. Given that the choice of demonstrated culture crucially depends on the adopted culture (since it induces preferences over demonstrated cultures), the parents have an incentive to actively engage in the culture formation process of their children.

However, the active engagement in the culture formation process of the child causes socialization costs. Besides the utility loss that parents experience when choosing a demonstrated culture that does not coincide with their adopted culture, they encounter cost of investments (especially time and effort) in their relative socialization success. The parental optimization problem in a strategic interactions setting is then the following. They choose a demonstrated culture and socialization success rates such as to maximize their life–time utility (i.e. the utility that they obtain out of their demonstrated culture and the utility related to the future demonstrated culture of their children), subject to the demonstrated cultures of the other adults and their socialization cost. Since the set of agents, from which a child socially learns in the socialization period is restricted to the adult generation only[10], we obtain through the parental optimization choices an *endogenous cultural formation of preferences* process.

[10]In this respect, we apply here an adoption of the 'direct vertical' and 'oblique' transmission framework that has been introduced into the economics literature by Bisin and Verdier (2000, 2001) and is a standard application in the literature on the economics of cultural transmission. The terminology originates from Cavalli-Sforza and Feldman (1981). Vertical transmission denotes the parental socialization contribution of a child and oblique transmission the socialization of the general adult society. These concepts are distinguished from 'horizontal transmission', which corresponds to socialization influences from members of the same generation.

Finally, we embed this framework in a society that is populated by distinct cultural groups such as to analyze the endogenous process of behavioral and cultural assimilation between the groups (*endogenous cultural distance*). The most important results of this analysis are summarized in the conclusion of this paper (section 5).

1.4 Outline

The remainder of this paper is organized as follows. Section 2 introduces in a general way the concept of cultural formation of preferences, while as section 3 endogenizes this process in a 'vertical' and 'oblique transmission' framework. Section 4 and 4.1 present the analysis on the endogenous behavioral and cultural assimilation of cultural groups, and section 5 concludes. All proofs of the propositions in this paper can be found in the appendix (which also contains an extension that did not find place in the main body of the text).

2 Cultural Formation of Preferences

Consider a society that is populated by a set of children, $c \in \mathcal{C}$, and adults, $a \in \mathcal{A}$ (both sets either finite or infinite). As far as the socialization process is concerned, we assume in the present paper that the children's social learning is only from the role–models of the adult members of the society, hence $\mathcal{A}$ does also constitute the (potential) socialization environment of the child. Let us assume for simplicity, that all adults have available the same set of feasible socio–economic choices, $X \in \mathbb{R}^M$, and let us denote with $n \in \mathcal{N} := \{1, \ldots, N\}$ the cultural values that are demonstrated through any choice therein[11].

Subject to its available options, the specific socio–economic choice of any adult $a \in \mathcal{A}$, $x_a \in X$, constitutes a demonstrated culture (the set of demonstrated cultural values), which will be represented by the mapping $\Phi^d : X \mapsto \Omega^d \subseteq \mathbb{R}^N$, where $\Phi^d(x_a) := \left(\phi^{1d}, \ldots, \phi^{Nd}\right)'(x_a)$ is the set of demonstrated cultural values, and where $\Omega^d = \Phi^d(X)$ denotes the *demonstrated culture set*. The latter does hence constitute the frame for the possible 'cultural behaviors' of the adults[12]. For ease of notation, we will for

[11] In general, the two dimensions M and N do not coincide. There can be interdependencies between different socio–economic actions both in terms of common (physical) constraints, and in terms of their joint meaning for one or more cultural values — and in the latter case, the set of cultural values for which two or more socio–economic actions have a joint meaning might not coincide with the sets of cultural values for which each of the socio–economic actions individually contributes meaning (the union of these sets might be larger than their intersection).

[12] One important issue to emphasize at this point is the fact that $\forall a \in \mathcal{A}$, the position of any $\Phi^d(x_a) \in \Omega^d$ is pinned down by the position of $x_a \in X$. Under this property, the scaling of the culture set is arbitrary (unless zero or infinite) and the mappings $b\Phi^d : X \mapsto b\Omega^d$,

subsequent use define, $\forall a \in \mathcal{A}$, $\Phi_a^d := \Phi^d(x_a)$.

As in the literature on the economics of cultural transmission, we assume that children are born without defined *culture*. Following and extending the basic idea of Bisin and Topa (2003) for the treatment of continuous traits[13], we will introduce a formal approach that represents the final culture of the children as the collective outcome of all socialization influences that the child has been exposed to during the socialization process. Thereby, we define a socialization influence as being constituted by a) the demonstrated culture, and b) the relative socialization success that this socialization agent could have achieved[14].

Let the demonstrated cultures of any of the adult persons $a \in \mathcal{A}$, i.e. the socialization environment of any child $c \in \mathcal{C}$, be $\Phi_a^d := (\phi_a^{1d}, \ldots, \phi_a^{Nd})' \in \Omega^d$, and for any cultural value $n \in \mathcal{N}$, let the socialization successes of these persons with the children be $\sigma_{ac}^n \in \mathbb{R}_+$. Then, the most general representation (at least up to the author's assessment) of the formation of continuous cultural values is the following[15].

Definition 1 (Formation of Cultural Values)**.** *The culture of a child,* $\Phi_c := (\phi_c^1, \ldots, \phi_c^N)'$*, is formed according to the* success–rate–weighted *demonstrated cultural values, i.e.* $\forall n \in \mathcal{N}$

$$\phi_c^n = \int_{a \in \mathcal{A}} \hat{\sigma}_{ac}^n \phi_a^{nd} da$$

where

$$\hat{\sigma}_{ac}^n := \frac{\sigma_{ac}^n}{\sigma_{ac}^n + \int_{a' \in \mathcal{A} \setminus \{a\}} \sigma_{a'c}^n da'}$$

denote the socialization success rates *of the respective persons.*

By Definition 1, we assume that the relative impact that any of the demonstrated cultural values of an adult has in the socialization process of

$b \in \mathbb{R} \setminus \{-\infty, 0, \infty\}$ would be equally suitable to represent the same demonstrated cultures. Notably, the properties of the mapping Φ^d are in general only sustained under linear scaling, but not under general affine transformations, since the demonstrated cultures could also contain entries that represent *relative* cultural values (especially in the case where two or more socio–economic actions that contribute meaning to different cultural values draw on the same resource, e.g. the time spent on different aspects of life).

Further note that if $b \in \mathbb{R}_{--}$, then this corresponds to a negative formulation of the original (formulation of the) cultural values.

[13]For an investigation of the issue of inheritance and formation of continuous traits in an empirical context see Otto et al. (1994).

[14]The logic of the socialization process that we present here could straightforwardly be extended to the socialization influence of peers as well as the society's institutions, like the legal and educational system or the media and marketing sector.

[15]We do deliberately not specify at this point whether the set of adult agents is finite or infinite, and use an integral representation since it is more general and encompasses also an eventual finite sum representation.

a child depends on the socialization success that the person can achieve for this demonstrated cultural value relative to the aggregated socialization successes that the socialization environment exerts on the child. Given the basic assumption that $\forall (a,c) \in \mathcal{A} \times \mathcal{C}$, $\sigma^n_{ac} \in \mathbb{R}_+$ we have that $\hat{\sigma}^n_{ac} \in [0,1]$ (they are really socialization success *rates*)[16] and then $\int_{a\in\mathcal{A}} \hat{\sigma}_{ac} da = 1$. Thus, the final cultural values of a child are convex combinations of all demonstrated cultural values it is exposed to. Intuitively, this means that the different demonstrated cultural values are mutually mediating in the culture formation process of the children. Also, the space where the final culture of any child can (in general) be located in, which we call the *culture set*, Ω, coincides with the convex hull of the demonstrated culture set, $\Omega = con\left(\Omega^d\right)$.

In the main body of this paper, we will for ease of the exposition refrain from a formal specification of potential inputs that lead to the achievement of socialization success. This stems from the fact that in general, there is a multitude of factors that determine the cognitive impact on the child (among which are notably the socialization interaction times as well as the effort and devotion a person invests into socializing a child) — and notably, these factors might differ for different cultural values[17]. Thus, one can not aim at including all potential factors into a single formal representation. The interested reader can, however, find in Appendix A.1.1 a proposal for a specification of the 'socialization success functions' that rests on generalizable determining factors, as well as a discussion of crucial properties of these functions.

To close the analysis in this subsection, let us consider the implications of the child's adoption, respectively internalization, of a specific culture after the completion of the socialization process. We just introduced this culture as being formed by the child's social learning from role–models (demonstrated cultures), realized through socio–economic choices out of the adults' feasible socio–economic choice sets. It seems then sensible to assume that, in the adult life–time of a child, its culture defines preferences over choices of socio–economic actions in the (same) feasible choice set and, thus, preferences over choices of demonstrated cultures in the (same) demonstrated culture set.

Assumption 1 (Preferences induced by Cultures)**.** *Let the adopted culture of an adult $a \in \mathcal{A}$ be $\left(\phi^1_a, \ldots, \phi^N_a\right) := \Phi_a \in \Omega$. Then, $\forall a \in A$,*

[16]Given this property, the socialization success rates could also be interpreted as the probabilities that a child adopts precisely the demonstrated culture of one specific adult as its own culture, i.e. $\forall (a,c) \in \mathcal{A} \times \mathcal{C}$, $P\left(\Phi_c = \Phi^d_a\right) = \hat{\sigma}_{ac}$. This constitutes the linking bridge between the treatment of discrete cultural traits or preferences and continuous cultural traits or preferences — since the way we constructed cultures makes both the probabilistic and convex mixing approach meaningful.

[17]In this respect, Panebianco (2009) shows the crucial dependence of the process of formation of inter–ethnic attitudes on whether the children use reciprocity and/or ethnocentrism as a 'filter' for evaluating the other groups' attitudes toward the own ethnic group.

(a) *cultures* $\Phi_a \in \Omega$ *induce complete and transitive preference relations* $\succ^{\Phi_a}$ *over demonstrated cultures* $\Phi_a^d \in \Omega^d$*, and*

(b) *the preferences* $\succ^{\Phi_a}$ *satisfy the following condition. Let* $\Phi_a^d, \Phi_a^{d'} \in \Omega^d$ *such that* $\text{sign}\left(\phi_a^n - \phi_a^{nd}\right) \in \left\{0, \text{sign}\left(\phi_a^n - \phi_a^{nd'}\right)\right\}$, $\forall n \in \mathcal{N}$*. Then* $\Phi_a^d \succ^{\Phi_a} \Phi_a^{d'} \Leftarrow \left|\phi_a^n - \phi_a^{nd}\right| \leq \left|\phi_a^n - \phi_a^{nd'}\right|$ $\forall n \in \mathcal{N}$*, and* $\left|\phi_a^n - \phi_a^{nd}\right| < \left|\phi_a^n - \phi_a^{nd'}\right|$ *for at least one* $n \in \mathcal{N}$*.*

We assume that adults prefer demonstrated cultural values that are as close in line as possible with the adopted cultural values. This view is in accordance with the *cognitive dissonance* theory of Festinger (1957). Both the adopted cultural values, as well as the demonstrated cultural values in the present text correspond to cognitive elements in Festinger's sense. Following his theory, people dislike dissonance between cognitive elements, the degree of which depends on the strength of the dissonance. Within the present text, the latter can be interpreted to correspond to the distance between the adopted and demonstrated cultural values.[18]

Finally, we have closed the circle between demonstrated cultures, cultures and preferences that constitutes the motivation to label the representation of the socialization process that this paper proposes as *cultural formation of preferences.*

3 Endogenous Cultural Formation of Preferences

The last section introduced in a general way the approach to the cultural formation of preferences that this paper proposes. In this section, we will

[18]In this context it follows that from the viewpoint of any adult the unique best element in the demonstrated culture set will be the one that exactly corresponds to the adopted culture (although this might not be contained in the demonstrated culture set if the latter is not convex), $\forall a \in \mathcal{A}$, $\Phi_a \succ^{\Phi_a} \Phi_a^d$, $\forall \Phi_a^d \in \Omega^d \backslash \{\Phi_a\}$. Giving definite statements about preference orderings of pairs of demonstrated cultures that do not involve the best element shall at this point only be pursued for the case where all pairs of demonstrated cultural values 'lie on the same side' of the respective adopted value or attitude (i.e. socio–economic choices that either reflect higher or lower intensities); and where this is adjoined by the case that the distances between the demonstrated cultural values and the adopted ones is (strictly) smaller for all entries. This stems from two reasons. First, as part of the socialization process, children might also adopt a preference with respect to the direction of the deviations of the demonstrated cultural values from their adopted ones — the possibility of which is though left unmodeled in the present paper. Second, it is in general not valid to assume that the adoption of a specific culture induces a preference relation of the form $\Phi_a^d \succ^{\Phi_a} \Phi_a^{d'} \Leftarrow d\left(\Phi_a, \Phi_a^d\right)_E < d\left(\Phi_a, \Phi_a^{d'}\right)_E$, where $\Phi_a^d, \Phi_a^{d'} \in \Omega^d$, i.e. that demonstrated cultures with a smaller (overall) Euclidean distance to the adopted culture are preferred. This assumption would restrict the analysis to preferences that weight all deviations between single cultural values and demonstrated cultural values equally, and would thus constitute a loss of generality.

propose a framework to endogenize this process as resulting out of optimal socialization decisions of parents. These decisions will be embedded in a framework of 'direct vertical transmission' (socialization inside the family) and 'oblique transmission' (socialization by the general social environment). This terminology stems from Cavalli-Sforza and Feldman (1981), which, in the context of the present paper, could straightforwardly be adopted to 'direct vertical *formation*' and 'oblique *formation*' (of preferences).

For ease of the exposition, we will assume here and throughout the rest of the paper that reproduction is asexual and that every adult has exactly one offspring. Under this assumption, we can denote with $\tilde{a} \in \tilde{\mathcal{A}}$ the children of the parents $a \in \mathcal{A}$ and the respective set of children, and with $\mathcal{A}_a := \mathcal{A} \backslash \{a\}$ the unrelated adult socialization environment of a child. Further, let us denote with $\hat{\sigma}_a^n := \hat{\sigma}_{a\tilde{a}}^n$, $\forall n \in \mathcal{N}$, the success rates that the parents have in the socialization of their child and with

$$\phi_{\mathcal{A}_a}^{nd} := \int_{a' \in \mathcal{A}_a} \frac{\sigma_{a'c}^n}{\sigma_{a'c}^n + \int_{a'' \in \mathcal{A}_a \backslash \{a'\}} \sigma_{a''c}^n da''} \Phi_a'^d da' \tag{1}$$

the relative–success–weighted demonstrated culture of the unrelated socialization environment. Then, we obtain the formation of cultural values (Definition 1) under direct vertical and oblique formation as

$$\phi_{\tilde{a}}^n = \phi_a^{nd} + (1 - \hat{\sigma}_a^n)(\phi_{\mathcal{A}_a}^{nd} - \phi_a^{nd}) \tag{2}$$

which we will also refer to as the *parental socialization technique*[19]. It embodies the view that parents set a cultural benchmark and can invest into their socialization success rates to countervail deviating cultural influences that the child is exposed to in the general (adult) social environment.

The motivation for parents to actively employ their socialization technique stems from the fact that the future demonstrated culture (which parents expect to depend on the future culture of the children; see below) has 'relevance' for their own life. This relevance can stem both from altruistic and egoistic reasons — and is established by socialization values (which could also be termed inter–generational cultural values).

Assumption 2 (Socialization Values)**.** $\exists \{1_s, \ldots, S_s\} \subset \mathcal{N}$, *such that* $\forall a \in \mathcal{A}$ *the socialization values,* $\Phi_a^s := \left(\phi_a^{1_s}, \ldots, \phi_a^{S_s}\right)'$, *induce complete and transitive preferences* $\succ^{\Phi_a^s}$ *over the future expected demonstrated culture of the own child,* $E_a\left(\Phi_{\tilde{a}}^d\right) \in \Omega^d$.

[19]Equation (2) is a generalization of the representation of the formation of continuous traits under vertical and oblique transmission in Bisin and Topa (2003). The latter was termed by Panebianco (2009) as 'Cavalli–Sforza Socialization Dynamics' (referring to the path–breaking work of Cavalli-Sforza and Feldman (1973, 1981) on issues related to different forms of cultural transmission and highlighting the inter–generational cultural dynamics that this kind of representation embodies).

Corollary 1 (Utility)**.** *Let Assumptions 1 and 2 hold. Then,* $\forall a \in \mathcal{A}$*,*

(a) the preferences $\succ^{\Phi_a}$ *can be represented by a utility function* $U^{\Phi_a} : \Omega^d \mapsto \mathbb{R}$*, where* $U^{\Phi_a}(\Phi_a) > U^{\Phi_a}\left(\Phi_a^d\right)$*,* $\forall \Phi_a^d \in \Omega^d \backslash \{\Phi_a\}$*, and*

(b) the preferences $\succ^{\Phi_a^s}$ *can be represented by an (expected) socialization utility function* $S^{\Phi_a^s} : \Omega^d \mapsto \mathbb{R}$*.*

Assuming that the 'own' utility function and the (expected) socialization utility function are additive separable, we may express the (expected) life–time utility of any adult $a \in \mathcal{A}$ as

$$U^{\Phi_a}\left(\Phi_a^d\right) + S^{\Phi_a^s}\left(E_a\left(\Phi_{\tilde{a}}^d\right)\right)$$

As far as the parental expectations of the future demonstrated culture of their children are concerned, we make here the standard assumption of the economics of cultural transmission literature on a form of parental myopia: Although parents have socialization values and thus take into account the impact of their chosen demonstrated culture on the future culture, and hence on the expected future demonstrated culture of their children, we assume that they do not realize this form of behavior changing impact that the socialization values impose on their children. Thus, parents expect their children to unrestrictedly maximize through their chosen demonstrated culture their 'own' utility, and hence $E_a\left(\Phi_{\tilde{a}}^d\right) \in \arg\max_{\Phi_a^d \in \Omega^d} U^{\Phi_a}\left(\Phi_a^d\right)$. Under the following convexity assumption (the compactness assumption will be needed later in this section), the culture set and demonstrated culture set coincide, $\Omega = \Omega^d$, which then guarantees by Corollary 1 (a) that $E_a\left(\Phi_{\tilde{a}}^d\right) = \Phi_{\tilde{a}}$, $\forall a \in \mathcal{A}$.

Assumption 3 (Convexity and Compactness)**.** *The demonstrated culture set* Ω^d *is convex and compact.*

Notably, any of the properties of Ω^d actually stems from the properties of the socio–economic choice set X under the mapping Φ^d. For guaranteeing compactness, it suffices that X is compact and the mapping Φ^d continuous. However, convexity of X together with continuity of Φ^d is not in general sufficient to assure convexity of Ω^d (but e.g. linearity of Φ^d would be).

Let us now discuss the (effort and devotion) costs associated with investments into the socialization successes. First, we will represent these costs by an indirect cost function of a choice of the parental socialization success rates. Second, we assume that the costs of a choice of some fixed $\hat{\sigma}_a^n$ depends also on how 'credible' the respective chosen demonstrated cultural value of the parents is. This is motivated from the fact that it is well known that persons that seem to be more satisfied with their chosen way of life have a more pronounced social learning impact (on children). Since the level of satisfaction (or utility) out of choices of demonstrated cultural values is determined by their position relative to the adopted cultural values (Assumption 1 (b)), we propose a cost function $C : [0,1]^n \times \Omega^d \times \Omega \mapsto \mathbb{R}_+$, $C\left(\hat{\Sigma}_a, \Phi_a^d, \Phi_a\right)$, where $\hat{\Sigma}_a := \left(\hat{\sigma}_a^1, \dots, \hat{\sigma}_a^N\right)'$.[20]

The parental optimization problem is it then to choose a demonstrated culture and the socialization success rates such as to maximize the expected life–time utility subject to socialization cost. We obtain, $\forall a \in \mathcal{A}$

$$\max_{\left(\Phi_a^d, \hat{\Sigma}_a\right) \in \Omega^d \times [0,1]^n} U^{\Phi_a}\left(\Phi_a^d\right) + S^{\Phi_a^s}\left(\Phi_{\tilde{a}}\right) - C\left(\hat{\Sigma}_a, \Phi_a^d, \Phi_a\right) \tag{3}$$

subject to the parental socialization technique (equations (2)). The optimization problems induce a game where parents have to decide about both their strategic interaction variables, Φ_a^d, and their non–strategic interaction variables, $\hat{\Sigma}_a$. As far as the latter are concerned, we assume that through their choice of the socialization success rates, the parents 'scale' the absolute socialization successes that all unrelated adults have with their children symmetrically, i.e. the relative socialization successes of equation (1) stay fixed.

We will introduce below two suitable equilibrium concepts, where the following notation will be useful. Denote $\Phi_{\mathcal{A}}^d := \left\{\Phi_a^d\right\}_{a \in \mathcal{A}}$ and $\hat{\Sigma}_{\mathcal{A}} := \left\{\hat{\Sigma}_a\right\}_{a \in \mathcal{A}}$; further let $\Omega_{\mathcal{A}}^d := \times_{a \in A} \Omega^d$ and $[0,1]_{\mathcal{A}}^n := \times_{a \in A} [0,1]^n$, as well as $\Phi_{-a}^d := \left\{\Phi_{a'}^d\right\}_{a' \in \mathcal{A}_a}$. Finally, denote the net (expected) life–time utility ('own' and socialization utility net of socialization cost) of any adult $a \in \mathcal{A}$ as $\mathcal{M}\left(\Phi_a^d, \hat{\Sigma}_a, \Phi_{-a}^d\right)$, and with $\Gamma(\cdot,\cdot,\cdot)$ a game where the first argument denotes the set of players, the second the utility of the individual players and the third the strategy space of the individual players.

Definition 2 (Equilibrium of a Period)**.**

(a) Call a tuple $\left(\Phi_{\mathcal{A}}^{d*}, \hat{\Sigma}_{\mathcal{A}}^*\right) \in \Omega_{\mathcal{A}}^d \times [0,1]_{\mathcal{A}}^n$ *a* weak equilibrium of a period (WEP) *if* $\Phi_{\mathcal{A}}^{d*}$ *is a pure strategy Nash Equilibrium of the game*

$$\Gamma\left(\mathcal{A}, \mathcal{M}\left(\Phi_a^d, \hat{\Sigma}_a^*, \Phi_{-a}^d\right), \Omega^d\right)$$

and $\forall a \in \mathcal{A}$*,* $\hat{\Sigma}_a^* \in \arg\max_{\hat{\Sigma}_a \in [0,1]^n} \mathcal{M}\left(\Phi_a^{d*}, \hat{\Sigma}_a, \Phi_{-a}^{d*}\right)$.

(b) Call a tuple $\left(\Phi_{\mathcal{A}}^{d**}, \hat{\Sigma}_{\mathcal{A}}^{**}\right) \in \Omega_{\mathcal{A}}^d \times [0,1]_{\mathcal{A}}^n$ *an* equilibrium of a period (EP) *if it is a pure strategy Nash Equilibrium of the game*

$$\Gamma\left(\mathcal{A}, \mathcal{M}\left(\Phi_a^d, \hat{\Sigma}_a, \Phi_{-a}^d\right), \Omega^d \times [0,1]^n\right).$$

[20] We silently made the assumption here that parents can choose all individual socialization success rates within the full unit interval. The deeper implications of this assumption are again discussed in Appendix A.1.1.

The WEP concept effectively decomposes the simultaneous parental optimization problems (3) into a sequential problem. Parents choose planned socialization success rates and then they choose their optimal demonstrated culture subject to this plan and the demonstrated cultures of the other adults. If the planned socialization success rates are then also optimal in the context of the demonstrated culture profile, a WEP is reached. For parents to play such strategies (that are not also EP strategies), we require a certain form of myopia, since they do not necessarily constitute local maximizers, let alone global maximizers, of the simultaneous optimization problems. But since the set of weak equilibria of a period obviously contains the set of equilibria of a period, any characterization of the chosen strategies of the first concept (which we will do in section 4) must also hold for the latter concept. Since the conditions for the existence of a WEP are weaker than that for the existence of an EP, the distinction of the two concepts is sensible.

Proposition 1 (Equilibrium Existence)**.** *Let Assumption 3 hold. Then,*

a) if the (hypothetical) optimization problems $\max_{\Phi_a^d \in \Omega^d} \mathcal{M}\left(\Phi_a^d, \hat{\Sigma}_a, \Phi_{-a}^d\right)$ *and* $\max_{\hat{\Sigma}_a \in [0,1]^n} \mathcal{M}\left(\Phi_a^d, \hat{\Sigma}_a, \Phi_{-a}^d\right)$ *are continuous and strictly concave* $\forall a \in \mathcal{A}$*, then a WEP exists.*

b) if the target functions of the optimization problems (3), $\max_{\left(\Phi_a^d, \hat{\Sigma}_a\right) \in \Omega^d \times [0,1]^n}$ $\mathcal{M}\left(\Phi_a^d, \hat{\Sigma}_a, \Phi_{-a}^d\right)$*, are continuous and quasi–concave* $\forall a \in \mathcal{A}$*, then an EP exists.*

Proof. In Appendix A.2.1.

At a first glance, the condition for the existence of a WEP does not look weaker than that for the existence of a EP — though it is on a closer look. The condition for existence of a WEP requires that U^{Φ_a}, $S^{\Phi_a^s}$ and $-C$ are jointly strictly concave in both $\hat{\Sigma}_a$ and Φ_a^d.[21] To discuss the conditions on the individual maps for quasi–concavity of problem (3) to hold, first note that the Hessian matrices of the parental socialization technique (2) are indefinite[22]. Thus, to guarantee at least quasi–concavity of the full optimization problem, the maps U^{Φ_a} and $-C$ must jointly be 'concave enough' (thus, effectively also strictly concave) to compensate for the eventual resulting non–concavities of $S^{\Phi_a^s}\left(\Phi_{\tilde{a}}\left(\Phi_a^d, \hat{\Sigma}_a\right)\right)$.[23]

[21]If $S^{\Phi_a^s}$ is (strictly) concave, then it is also (strictly) concave in both $\hat{\Sigma}_a$ and Φ_a^d. This holds since when fixing the entries of one vector, the parental socialization technique (2) is linear in the entries of the other vector.

[22]$\forall a \in \mathcal{A}$, $\forall n \in \mathcal{N}$, $\left|H_{\phi_{\tilde{a}}^n\left(\phi_a^{n\,d}, \hat{\sigma}_a^n\right)}\right| = -1$.

[23]In the following section, we will introduce a corresponding assumption for twice continuously differentiable functions.

The aim of the present section was to introduce a general framework to endogenize the cultural formation of preferences process and to provide for conditions for a (weak) equilibrium of a period to exist. For a characterization of the choices in a (weak) equilibrium of a period, as well as an analysis of the inter–generational evolution of the cultures of the agents, we will have to further specify the model. This is the subject of the following section.

4 Endogenous Cultural Distance

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. We define a cultural group as the collection of families for which it holds that the members of the parental generation have equal culture. We will provide for a characterization of weak equilibria of a period (hence also for equilibria of a period), a discussion of comparative statics effects of the model framework (both in the present section 4), as well as we will represent the dynamics of the cultural distance between the two cultural groups in section 4.1 (in general, we define the cultural distance as the Euclidean distance between the cultures of the groups). To ease the analytical tractability, the exposition will be based on a number of simplifications compared to the general framework of section 3.

Consider an overlapping generations society that is (initially) populated by two distinct cultural groups, indexed $G \in \{L, H\}$, with the population shares denoted q_G. In the present setting, it will be convenient to index the (sets of) adults and children as $g \in \mathcal{G}$ and $\tilde{g} \in \tilde{\mathcal{G}}$, and to denote $-G := \{L, H\} \backslash \{G\}$. In any given period, all adult members of a cultural group have equal culture (we will introduce below conditions that assure this property), which we reduce for the present analysis to consist of only one 'own' cultural value and one socialization value, i.e. $\forall G \in \{L, H\}$, $\forall g \in \mathcal{G}$, $\Phi_g = \Phi_G := (\phi_G, \phi_G^s)'$.

Since parents are myopic and do not consider the behavior changing impact of their children's adopted socialization value(s), it is immediate that their children's future demonstrated socialization value(s) do also not enter into their socialization utility function (since if they would enter, we could also not sensibly maintain the parental myopia assumption), so that $S^{\phi_G^s}(\Phi_{\tilde{g}}) = s^{\phi_G^s}(\phi_{\tilde{g}})$. Since parents do effectively not care for the formation of their children's socialization values, we assume for simplicity that they are inter–temporarily fixed.

Continuing this line of assumptions, we do in the present analysis also not consider the 'own' utility contribution of the parents' chosen displayed socialization values, which we denote ϕ_g^{sd}. In this context, we assume for convenience that total 'own' utility is additive separable in the utility out of the endogenous demonstrated cultural value and the utility out of the demonstrated socialization value, $U^{\Phi_G}\left(\Phi_g^d\right) = u^{\phi_G}\left(\phi_g^d\right) + u^{\phi_G^s}\left(\phi_g^{sd}\right)$, and consider subsequently only the first component[24].

Finally, two additional specifications shall be introduced. The first concerns the socialization values. Within the present analysis, it is interesting to consider different classes of socialization values, which we specify by the (usual) strength of the socialization value, ϕ_G^s, and the class–specific target for the future demonstrated 'own' cultural value of the children (which accords to the future culture of the children under the parental myopia), $\hat{\phi}_G$.[25] The according complete notation for the (myopic) socialization utility is then $s^{\phi_G^s}\left(\phi_{\tilde{g}}, \hat{\phi}_G\right)$.

The second specification concerns the role that the position of the chosen demonstrated 'own' cultural value relative to the adopted one plays in the indirect socialization cost functions. Remember from section 3 that we motivated their relevance as stemming from their determination of the level of satisfaction that an adult can achieve through its choice of demonstrated culture. It is then immediate to represent the indirect cost (of the one–dimensional case) as a function $c : [0,1] \times \mathbb{R} \mapsto \mathbb{R}_+$, $c\left(\hat{\sigma}_g, u^{\phi_G}\left(\phi_g^d\right)\right)$.

Denoting with ω^d the choice set of demonstrated 'own' cultural values, we finally obtain the (reduced) parental optimization problems (3) as

$$\max_{\left(\phi_g^d, \hat{\sigma}_g\right) \in\, \omega^d \times [0,1]} u^{\phi_G}\left(\phi_g^d\right) + s^{\phi_G^s}\left(\phi_{\tilde{g}}, \hat{\phi}_G\right) - c\left(\hat{\sigma}_g, u^{\phi_G}\left(\phi_g^d\right)\right) \qquad (4)$$

subject to the parental socialization techniques (2).

Given the fixed socialization value (and the disregarding of the demonstrated socialization values of the adults), we will concentrate the analysis below on the formation and evolution of the endogenous cultural value. In this light, we will (in an inconsistency with the previous representation of the cultures of the adults) subsequently refer to the single endogenous cultural value as the *culture* of a group (unless otherwise noted). Accordingly, we will denote as *cultural distance* the Euclidean distance between the endogenous cultural values of the two groups[26], $\Delta^\phi := |\phi_L - \phi_H|$.

---

[24] Since parents do not consider the impact of their demonstrated socialization value on the future demonstrated socialization value of their children, they can unrestrictedly maximize the second utility component, i.e. choose $\phi_g^{sd} = \phi_G^s$ — we leave however undiscussed at this point the foundations and implications of such a choice.

[25] We will consider in the subsequent analysis both a characterization of a WEP (restricted to cases where all parents of a cultural groups make identical decisions; see below) and comparative statics under a general class of socialization values in the present section, as well as we will consider the dynamics of the cultures for the case where the socialization target coincides with the cultures of the parents in section 4.1. Within the present framework, there arise two other (but unconsidered) natural variants for group–specific 'socialization targets', namely an exogenously fixed target, and the representative displayed culture of the adult members of a cultural group.

[26] Obviously, in the present context, the true cultural distance between the two groups would be $d\left(\Phi_L, \Phi_H\right)_E$.

Assumption 4 (Curvature I)**.** $\forall G \in \{L, H\}$, $\forall g \in \mathcal{G}$,

(a) the maps u^{ϕ_G} and $s^{\phi^s_G}$ are continuous[27] *and concave and $-c$ is continuous and concave in both arguments, and*

(b) either $s^{\phi^s_G}$ is strictly concave, and/or u^{ϕ_G} is strictly concave and $-c$ is strictly concave in $\hat{\sigma}_g$, and/or $-c$ is strictly concave in both arguments.

It is immediate that under Assumption 4, the conditions for the existence of a WEP are satisfied, since then both (hypothetical) optimization problems in Proposition 1 (a) are strictly concave. At this point, we introduce two further simplifications that are crucial for the subsequent analysis: First, we assume that the relative aggregated socialization successes of the unrelated members of both cultural groups in the socialization process of any child coincide with the ratio of the population shares of the groups, i.e. $\forall G \in \{L, H\}$, $\forall \tilde{g} \in \tilde{\mathcal{G}}$

$$(1 - q_G) \int_{g' \in \mathcal{G}_g} \sigma_{g'\tilde{g}} \, dg' = q_G \int_{-g \in -\mathcal{G}} \sigma_{-g\tilde{g}} \, d{-g} \tag{5}$$

where $\mathcal{G}_g := \mathcal{G} \backslash \{g\}$. Second, we consider only weak equilibria of a period, in which all parents of a cultural group make identical decisions[28]. We will call such a WEP a *homogeneous weak equilibrium of a period* (HWEP).

Proposition 2 (Existence of HWEP)**.** *If Assumption 1—4 and condition (5) hold, then a HWEP exists.*

Proof. In Appendix A.2.2.

Since under the set–up of the present section, the optimization problems of the adult members depend on their own culture, the socialization target culture and the population shares, we will denote the group–homogeneous parental decisions in a HWEP as a tuple

$$\left\{\phi_G^{d^*}\left(\phi_L, \phi_H, \phi^s_L, \phi^s_H, \hat{\phi}_L, \hat{\phi}_H, q_H\right), \hat{\sigma}^*_G\left(\phi_L, \phi_H, \phi^s_L, \phi^s_H, \hat{\phi}_L, \hat{\phi}_H, q_H\right)\right\}_{G \in \{L,H\}}.$$

Using these in the parental socialization technique (2), we obtain the endogenous future homogeneous culture of the children of both groups, $\forall G \in \{L, H\}$

$$\phi^*_{\tilde{G}}(P) := \phi_G^{d^*}(P) - \left(\phi_G^{d^*}(P) - \phi_{-G}^{d^*}(P)\right)(1 - \sigma^*_G(P))(1 - q_G)$$

[27]Note here that under Assumption 3, the utility functions would also be guaranteed to be continuous (since then the preferences are defined on a convex set).

[28]Notably, under the last two simplifying assumptions, it follows that all children of a cultural group adopt the same culture, so that the population shares of the two distinct groups are constant over generations.

where $P := \{\phi_L, \phi_H, \phi^s_L, \phi^s_H, \hat{\phi}_L, \hat{\phi}_H, q_H\}$. Obviously, these HWEP maps do also endogenously determine the future cultural distance between the two groups as

$$\begin{aligned}\tilde{\Delta}^{\phi^*}(P) :&= \left|\phi^*_{\tilde{L}}(P) - \phi^*_{\tilde{H}}(P)\right| \\ &= \left|\left(\phi^{d^*}_L(P) - \phi^{d^*}_H(P)\right)\left(\hat{\sigma}^*_L(P) q_H + \hat{\sigma}^*_H(P)(1-q_H)\right)\right|\end{aligned}$$

In the rest of this section, we will introduce additional assumptions that allow for a characterization both of a HWEP — and thus also of a homogeneous equilibrium of a period (HEP) — and of the comparative statics of the cultural system.

Assumption 5 (Slope)**.** *$\forall G \in \{L, H\}$, $\forall g \in \mathcal{G}$,*

(a) the mappings u^{ϕ_G} and $s^{\phi^s_G}$ are $\mathcal{C}^1$, and $s^{\phi^s_G}$ is strictly decreasing on both sides of $\hat{\phi}_G$,

(b) the map c is $\mathcal{C}^1$, strictly increasing in the first argument $\forall \hat{\sigma}_g \in (0,1]$, where $\frac{\partial c(0,\cdot)}{\partial \hat{\sigma}_g} = 0$, and decreasing in the second argument.

Notably, under Assumption 1 (b), u^{ϕ_G} is also strictly decreasing on both sides of ϕ_G. Together with Assumption 5 (a), this implies that $\frac{\partial u^{\phi_G}(\phi_G)}{\partial \phi^d_g} = 0$, and furthermore $\frac{\partial s^{\phi^s_G}(\hat{\phi}_G, \hat{\phi}_G)}{\partial \phi_{\tilde{g}}} = 0$. For the following Proposition, let us denote with ϕ_{G_0} the initial cultures of both groups $G \in \{L, H\}$, and denote $\omega^d := \left[\underline{\phi}, \overline{\phi}\right]$.

Proposition 3 (Characterization of a HWEP)**.** *Let Assumptions 1–5 and condition (5) be satisfied. Then, if $\phi_{H_0} > \phi_{L_0}$ and if $\hat{\phi}_L \leq \phi_L$ and $\hat{\phi}_H \geq \phi_H$ holds in any period*[29]*, the following properties are satisfied in any evolving HWEP*

1. Case $\phi_H > \phi_L$:

(a) $\hat{\sigma}^*_G \in (0,1]$, $\forall G \in \{L, H\}$

(b) $\phi^{d^*}_H \geq \phi_H > \phi_L \geq \phi^{d^*}_L$ (with equalities iff $\phi_H = \overline{\phi}$ or $\phi_L = \underline{\phi}$)

(c) $\hat{\phi}_H > \phi^*_{\tilde{H}} > \phi^*_{\tilde{L}} > \hat{\phi}_L$

[29] Restricting the analysis to cases where the initial cultures of the two groups are unequal, we can without loss of generality assume that $\phi_{H_0} > \phi_{L_0}$. The condition $\hat{\phi}_L \leq \phi_L$ and $\hat{\phi}_H \geq \phi_H$ holds per definition in the case of the inter–generational cultural closeness socialization motivation (section 4.1), since then $\hat{\phi}_G \equiv \phi_G$ in any period.

Further note for completeness that we consider here only classes of socialization values that are either independent of individual parental decisions of demonstrated cultural values by construction, or where parents believe this to hold (i.e. they do not internalize the eventual effect of their decisions on the socialization target).

2. *Case* $\phi_H = \phi_L$*:*

   (a) $\hat{\sigma}^*_G \in [0,1]$, $\forall G \in \{L, H\}$

   (b) $\phi^{d^*}_H \geq \phi_H = \phi_L \geq \phi^{d^*}_L$

   (c) $\hat{\phi}_H \geq \phi^*_{\tilde{H}} \geq \phi^*_{\tilde{H}} \geq \hat{\phi}_L$

Proof. In Appendix A.2.3.

By the results of Proposition 3, 1., it follows that if the cultural system starts with initially different cultures, then in any HWEP (as long as the cultural system has not converged to a point where the cultures of both groups are homogeneous), parents of both groups will, if possible, dis–integrate behaviorally (i.e. they choose a more 'radical' demonstrated culture relative to the demonstrated culture of the members of the other groups, than the choice of their true culture would be), which implies $\phi^{d^*}_H - \phi^{d^*}_L =: \Delta^{\phi^{d^*}} > \Delta^{\phi}$. This comes in attempt to countervail the 'negative' cultural influence that the adult members of the other cultural group exert on the child. Of course, such a deviation from the parents' utility maximizing demonstrated cultural value ($\phi^d_g = \phi_G$) must be accompanied with strictly positive investments into the socialization success rate. Since parents do not perceive a socialization utility loss for marginal deviations of the child's culture (respectively expected demonstrated culture) from the socialization target cultures, the future cultures of the children of both cultural groups will always lie strictly in the interior of the cultural space that is formed by the target cultures.

Furthermore, given the strictly positive socialization success rates that the parents choose in any HWEP, the final culture of their children is closer to the target culture than the (relative–success–weighted) representative demonstrated culture of the unrelated adult environment is (which is equal for all parents and children under condition (5)). This means that the relative positions of the cultures of the two groups (in terms of a lower or higher endogenous cultural value) are preserved over generations. Thus, although the cultures of both groups are subject to endogenous dynamical change, the 'labels' of the two cultural groups are dynamically persistent. Given the additional results of Proposition 3, 2., this result holds weakly even when the cultures of both groups should have converged to a homogeneous culture.

To guarantee the existence of a HEP, and for the subsequent comparative statics results, we have to apply stronger assumptions (in analogy to previous notation, we denote here with ϕ_{-g} the tuple of demonstrated cultures of the set of adults $-g := \mathcal{G}_g \cup -\mathcal{G}$).

Assumption 6 (Curvature II)**.** $\forall G \in \{L, H\}$,

(a) the maps u^{ϕ_G}, $s^{\phi^s_G}$ *and* c *are* $\mathcal{C}^2$*, and* $\frac{\partial^2 \mathcal{M}(\phi^d_g, \hat{\sigma}_g, \phi^d_{-g})}{\partial {\phi^d_g}^2} \frac{\partial^2 \mathcal{M}(\phi^d_g, \hat{\sigma}_g, \phi^d_{-g})}{\partial \hat{\sigma}^2_g} \geq \left(\frac{\partial^2 \mathcal{M}(\phi^d_g, \hat{\sigma}_g, \phi^d_{-g})}{\partial \phi^d_g \, \partial \hat{\sigma}_g} \right)^2$*, and*

(b) *the marginal cost of a deviation of the future (demonstrated) culture of the children from the target culture is increasing in the socialization value,* $\mathrm{sign}(\hat{\phi}_G - \phi_{\tilde{G}})\frac{\partial^2 s^{\phi^s_{\tilde{G}}}(\phi_{\tilde{G}},\phi^s_{\tilde{G}})}{\partial \phi_{\tilde{G}}\, \partial \phi^s_{G}} \geq 0.$

Proposition 4 (Existence of HEP)**.** *If Assumptions 1–4, and 6 (a) as well as condition (5) hold, then a HEP exists.*

Proof. In Appendix A.2.4.

In analogy to previous notation, we will denote the group–homogeneous parental decisions in a HEP as a tuple

$$\left\{\phi^{d^{**}}_G\left(\phi_L, \phi_H, \phi^s_L, \phi^s_H, \hat{\phi}_L, \hat{\phi}_H, q_H\right), \hat{\sigma}^{**}_G\left(\phi_L, \phi_H, \phi^s_L, \phi^s_H, \hat{\phi}_L, \hat{\phi}_H, q_H\right)\right\}_{G\in\{L,H\}}.$$

Proposition 5 (Comparative Statics)**.** *Let Assumptions 1–6 be satisfied, and let us distinguish the cultural groups as $G = L/H$. Then, in any HEP where $\phi_L \neq \phi_H$,*[30] $\exists\ (b_{G1}(P), b_{G2}(P)) < / > (0,0)$*, where* $b_{G1}(P)b_{G2}(P) \geq 1$*, such that if* $\forall\ G \in \{L, H\}$

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{\partial^2 \mathcal{M}\left(\phi^{d^{**}}_G, \hat{\sigma}^{**}_G, \phi^{d^{**}}_{-g}\right)}{\partial \phi^d_g\, \partial \hat{\sigma}_g} < / > \begin{pmatrix} b_{G1}(P)\frac{\partial^2 \mathcal{M}\left(\phi^{d^{**}}_G, \hat{\sigma}^{**}_G, \phi^{d^{**}}_{-g}\right)}{\partial {\phi^d_g}^2} \\ b_{G2}(P)\frac{\partial^2 \mathcal{M}\left(\phi^{d^{**}}_G, \hat{\sigma}^{**}_G, \phi^{d^{**}}_{-g}\right)}{\partial \hat{\sigma}^2_g} \end{pmatrix}$$

the following comparative statics results hold[31]

(a)

$$\mathrm{sign}\begin{pmatrix} \frac{\partial \phi^{d^{**}}_G}{\partial \phi_G} & \frac{\partial \phi^{d^{**}}_G}{\partial \phi_{-G}} & \frac{\partial \phi^{d^{**}}_G}{\partial \phi^s_G} & \frac{\partial \phi^{d^{**}}_G}{\partial \phi^s_{-G}} & \frac{\partial \phi^{d^{**}}_G}{\partial \hat{\phi}_G} & \frac{\partial \phi^{d^{**}}_G}{\partial \hat{\phi}_{-G}} \\ \frac{\partial \sigma^{**}_G}{\partial \phi_G} & \frac{\partial \sigma^{**}_G}{\partial \phi_{-G}} & \frac{\partial \sigma^{**}_G}{\partial \phi^s_G} & \frac{\partial \sigma^{**}_G}{\partial \phi^s_{-G}} & \frac{\partial \sigma^{**}_G}{\partial \hat{\phi}_G} & \frac{\partial \sigma^{**}_G}{\partial \hat{\phi}_{-G}} \end{pmatrix} = \begin{pmatrix} +/+ & -/- & -/+ & -/+ & +/+ & -/- \\ -/+ & +/- & +/+ & +/+ & -/+ & +/- \end{pmatrix}$$

(b)

$$\left(\mathrm{sign}\left(\frac{\partial \phi^{d^{**}}_L}{\partial q_H}\right), \mathrm{sign}\left(\frac{\partial \sigma^{**}_L}{\partial q_H}\right)\right) \in \{(-,+),(+,-),(0,0)\}$$

$$\left(\mathrm{sign}\left(\frac{\partial \phi^{d^{**}}_H}{\partial q_H}\right), \mathrm{sign}\left(\frac{\partial \sigma^{**}_H}{\partial q_H}\right)\right) \in \{(+,+),(-,-),(0,0)\}$$

Proof. In Appendix A.2.5.

[30] If $\phi_L = \phi_H$ is a steady state as in section 4.1, then marginal changes in the states and the parameters will cause no changes in the optimal choices of the parents.

[31] Note that in case $\hat{\phi}_G \equiv \phi_G$ (section 4.1) the comparative statics effects of the first and third column as well as of the second and fourth column have to be combined, which leaves the signs unchanged.

Proposition 5 requires that for both cultural groups, in any HEP, the two socialization instruments (decision variables) must not be too strong substitutes in the parental socialization problem (4) — the strength of which is determined by an expression that weights the concavities of problem (4) in the two socialization instruments. If this condition together with Assumptions 1–6 are fulfilled, then the parents will always alter the use of the two instruments in a symmetric way. This means that an increase of behavioral dis–integration will be adjoined by an increase of the socialization success rate. To understand this behavior, one has to be aware of the combined impact of two different effects. The first and direct effect is caused by a change in the own culture or 'own' parameters of a cultural group. The second, strategic interaction, effect is caused by the mutual reactions of parents of both cultural groups to changes in the demonstrated culture of the other group.

If the condition in Proposition 5 is fulfilled for one cultural group $G \in \{L, H\}$, then the parents of that group would react to the following changes by increasing both socialization instruments (the direct effects): An increase of the distance of the own culture to the (demonstrated) culture of the other group (since the own utility loss for 'more radical' cultural behavior is decreased); an increase of the socialization target culture to the (demonstrated) culture of the other group (since the increased distance of the future culture of the children to the socialization target culture decreases expected utility)[32]; an increase of the socialization motivation intensity (since equal distances of the culture of the children to the target culture cause lower expected utility); and an increase of the distance of the demonstrated culture of the other group to the own socialization target culture (which increases the distance of the future culture of the children to the socialization target culture).

The central question now is whether Proposition 5 is also fulfilled for the other cultural group $-G$. This then would imply that the parents of that group would react to an increase of the distance of the demonstrated culture of group G to their socialization target culture by increasing the distance of their demonstrated culture to the (demonstrated) culture of group G. In this case, both the direct effect of a change in the culture or parameters of a group and the strategic interaction effect work in the same direction and the overall effect of changes in the cultures or parameters is unambiguous. Since the direct effects as described above for cultural group G are also valid for group $-G$, the remaining comparative statics effects of Proposition 5 (a) for changes in the culture or parameters of the other group ($-G$) are due to the analogous reasoning.

The only parameter–change that has a direct impact on the optimal decisions of both cultural groups concerns the population share $q_H = 1 - q_L$. An

[32] In the case $\hat{\phi}_G \equiv \phi_G$ the first two effects are combined.

increase (decrease) of the own population share improves (dis–improves) the cultural composition of the public social space for the parents of a cultural group. This means that it becomes more (less) attractive as a substitute for own socialization investments and parents would, under the condition of Proposition 5, react by decreasing (increasing) them. Nevertheless, the strategic interaction effect works in exactly the other direction since parents would react to less (more) distant demonstrated cultures of the other cultural group by decreasing (increasing) the distance of their own demonstrated culture to the demonstrated culture of the other group. The relative strength of the direct and strategic interaction effects will then determine the direction of the overall changes in the use of the socialization instruments (which could also be zero), which illustrates Proposition 5 (b).

The following section specifies the properties of the model presented so far under a class of socialization values that is related to the parental desire for *inter–generational cultural closeness*. The analysis is extended to derive properties of the dynamics of the cultural system.

4.1 Inter–generational Cultural Closeness

It is an undisputed fact that family is one of the most central aspects in human life. This section introduces a class of socialization values that is based on intra–family considerations of the parents. Specifically, this class represents a desire of the parents for the future culture of their children to be close to their own culture — which hence serves as the target culture in the socialization process. Subsequently, we will shortly introduce three justifications for this assumption.

The first justification is based on a special form of parental altruism called 'imperfect empathy'. This concept has been introduced into the economics literature by Bisin and Verdier (2000, 2001). Parents are altruistic to their children and want to maximize the utility out of the children's future socio–economic choices, based on the expectations of their culture. Nevertheless, parents can assess this utility only through the filter of their own preferences (culture). In the present framework, this introduces a desire for cultural closeness, since the parents are myopic and expect their children to choose a life–style that accords to their culture. A second interpretation of family related values is what we call 'family cohesion', and is based on social interactions within the family, thus on the demonstrated cultures of the generations. If the demonstrated cultures of parents (in their third life–period) and children (in their adult–period) are more distant, then this causes more frictions, misunderstandings and (mutual) disappointment. Again given the parental myopia, and under the assumption that the parents expect to set, in their third life–period, a demonstrated culture that equals their true culture[33], this then justifies the assumption of the parental desire for inter–generational cultural closeness. The third, and purely philosophical, interpretation is based on the assumption that parents simply have an intrinsic desire for their children to develop a personality, hence culture, that is similar to their own one (e.g. if they view the life of their children as a continuation of their own life)[34].

Given these motivations for a desire for inter–generational cultural closeness, we have that $\hat{\phi}_G \equiv \phi_G$, and the expected socialization utility function is $s^{\phi^s_G}(\phi_{\tilde{g}}, \phi_G)$.[35] Using this to specify optimization problem (4), we have the following characterization of a HWEP under the parental desire for inter–generational cultural closeness.

Proposition 6 (Characterization of a HWEP)**.** *Let Assumptions 1–5 and condition 5 be satisfied. Then if $\phi_{H_0} > \phi_{L_0}$, the following properties are satisfied in any evolving HWEP*

1. *Case $\phi_H > \phi_L$*
   (a) $\hat{\sigma}^*_G \in (0, 1)$, $\forall G \in \{L, H\}$
   (b) $\phi^{d^*}_H \geq \phi_H > \phi^*_{\tilde{H}} > \phi^*_{\tilde{L}} > \phi_L \geq \phi^{d^*}_L$
   (with equalities iff $\phi_H = \overline{\phi}$ or $\phi_L = \underline{\phi}$)
2. *Case $\phi_L = \phi_H$*
   (a) $\hat{\sigma}^*_G = 0$, $\forall G \in \{L, H\}$
   (b) $\phi^{d^*}_H = \phi_H = \phi^*_{\tilde{H}} = \phi^*_{\tilde{L}} = \phi_L = \phi^{d^*}_L$

Proof. In Appendix A.2.6.

[33] In the third life–period, adults have no more need to engage in socialization, and thus can unrestrictedly maximize their utility out of socio–economic choices.

[34] Thus, the latter justification for inter–generational cultural closeness class is directly related to the future culture of the children (while as the other justifications were only indirectly related given the parental myopia). We nevertheless refrained from introducing a separate notation for such cases in section 3, given their exceptionality.

[35] In the present case, ϕ^s_G can be considered as a composed value out of the usual time–discount rate (since the relevant socio–economic actions of the children take place one period after the parents take the socialization actions), and a value that represents the strength of the parental desire for inter–generational cultural closeness.

In the case of imperfect empathy, it is immediate to specify $s^{\phi^s_G}(\phi_{\tilde{g}}, \phi_G) := \phi^s_G u^{\phi_G}(\phi_{\tilde{g}})$, where ϕ^s_G then corresponds to the (discounted) 'degree of parental altruism' (i.e. the valuation that parents put on the future utility of their children relative to their 'own' utility). The formulations in the literature on the economics of cultural transmission, notably Bisin and Verdier (2000, 2001), Bisin and Topa (2003) and Panebianco (2009), correspond to the special case $\phi^s_G = 1$. Given imperfect empathy, the utility loss that parents perceive if their children do not exactly adopt the parental culture is then $V^{\phi^s_G}(\phi_{\tilde{g}}, \phi_G) := \phi^s_G \left(u^{\phi_G}(\phi_G) - u^{\phi_G}(\phi_{\tilde{g}})\right)$.

Compared to the general results of Proposition 3, in case of the class of inter–generational cultural closeness socialization values, where the socialization target coincides with the culture of the parents, they would never choose a combination of socio–economic actions and socialization success rate such that their children adopt exactly the parental culture. This comes from the fact that if the future culture of the children only marginally deviates from the parents' culture, this is not perceived as costly. As a result, the final cultures of the children of both cultural groups lie in the interior of the interval that is constituted by the cultures of the parental generation, $\forall G \in \{L, H\}$, $\phi^*_{\tilde{G}} \in (\phi_L, \phi_H)$. This then implies that over generations, the cultures of the groups assimilate (but not completely) $\Delta^{\tilde{\phi}^*} < \Delta^{\phi}$. Asymptotically, though, the cultures converge to a homogeneous culture, where $\phi_H = \phi_L$, and parents have no more need to engage into active socialization, so that this homogeneous culture is also a rest point. While as the exact location of the homogeneous steady state depends on the initial location of the cultures (and the socialization values and population shares), the corresponding cultural distance of zero in a homogeneous steady state is globally asymptotically stable.

This result will be stated more formally in a Corollary. First, consider the present overlapping generations model in the continuous time limit[36]. Let the paths of cultures that evolve out of the parental problems (4) under the socialization utility function $s^{\phi^s_G}(\phi_{\tilde{g}}, \phi_G)$ and under the conditions of a HWEP be denoted $\phi_H(t, \phi_{H_0})$ and $\phi_L(t, \phi_{L_0})$, where t is the time–index. The corresponding path for the endogenous cultural distance is denoted $\Delta^{\phi}(t, \Delta^{\phi}_0) := \phi_H(t, \phi_{H_0}) - \phi_L(t, \phi_{L_0})$, where $\Delta^{\phi}_0 := \phi_{H_0} - \phi_{L_0}$.

Corollary 2 (Dynamics of Cultural Distance)**.** *If Assumptions 1–5 are fulfilled, then* $\lim_{t\to\infty} \Delta^{\phi}(t, \Delta^{\phi}_0) = 0$, $\forall \Delta^{\phi}_0 \in \left[0, \overline{\phi} - \underline{\phi}\right]$. *Moreover,* $\forall G \in \{L, H\}$, $\lim_{t\to\infty} \phi_G(t, \phi_{G_0}) = \phi^* \in (\phi_{L_0}, \phi_{H_0})$.

Proof. By Proposition 6, 1.(b), the dynamics of the cultural distance is a contraction mapping and the system converges to a homogeneous culture for any initial cultural distance. By Proposition 6, 2., any homogeneous culture is a rest point. The last result follows directly from Proposition 6. □

By Corollary 2, the social(ization) interactions of the two cultural groups induce a dynamic cultural assimilation process, which results in the creation of a new homogeneous equilibrium culture that can be interpreted as a mixture of the two initial cultures. This theoretical result corresponds to the 'melting pot' theory of cultural assimilation (see Han, 2006, p. 32).

[36] We take the derivation of the continuous time limit from Bisin and Verdier (2001): In an OLG economy where the agents live Δ units of time and have children $1-h$ units of time after birth, one obtains the continuous time limit by taking the limit for $\Delta, h \to 0$ with $\frac{h}{\Delta} \to 0$.

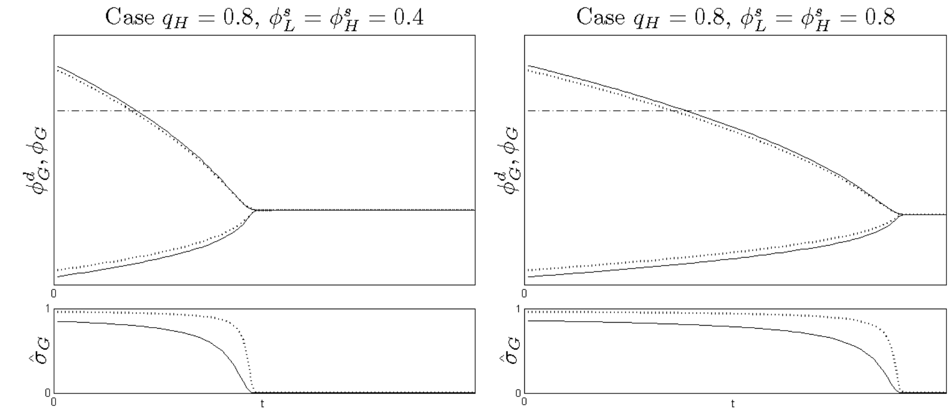

Figure 1: Dynamics of the Cultural System

To close the analysis of the cultural system under the inter–generational cultural closeness socialization motivation, we illustrate the analytical results by means of a numerical simulation[37]. Figure 1 shows the dynamics of the cultural system for two majority/minority cases. In the upper graph of each case, the solid lines represent the demonstrated culture of the groups and the dotted lines represent the path of the cultures of the groups (and the dash–dotted line locates the population–share weighted convex combination of the initial cultures; this would equate to the cultural steady state if parents of both cultural groups would not invest into the socialization instruments). In both graphs, the 'low value' cultural group is the minority, whose path of the socialization success rate is represented by the dotted lines in the lower graph of each case.

In both parts of Figure 1, one can see that parents choose a behavioral dis–integration as well as strictly positive socialization success rates as long as the cultural distance between the groups is strictly positive. Nevertheless, parents never choose high enough investments into the socialization instruments to assure that the cultural distance between the two groups would be eventually non–decreasing, which leads to an inter–temporal assimilation process toward a homogeneous steady state culture (Proposition 6, and Corollary 2). Since (if the conditions of Proposition 5 hold) the demonstrated cultural distance as well as the socialization success rates are strictly declining for decreases in the cultural distance (this follows directly from Proposition 5 when $\phi_G \equiv \hat{\phi}_G$), they are also strictly declining throughout

[37]For the numerical simulation, we used the following specifications: $u^{\phi_G}\left(\phi_g^d\right) = -\left|\phi_G - \phi_G^d\right|^2$, $s^{\phi_G^s}\left(\phi_{\tilde{g}}, \phi_G\right) = -\phi_G^s \left|\phi_G - \phi_{\tilde{g}}\right|^2$, $c\left(\hat{\sigma}_g, u^{\phi_G}\left(\phi_g^d\right)\right) = \hat{\sigma}_g^2 \left(1 - u^{\phi_G}\left(\phi_g^d\right)/20\right)$. It is immediate that these functions satisfy all relevant Assumptions of this section. Furthermore, we used an initial cultural distance of 28 units, and the total length of the time–axis in all graphs corresponds to 100 periods. Nevertheless, the choice of the initial cultural distance and the resulting length of the convergence path could be arbitrarily changed, so that they are not indicated in the graphs.

the transitory path, i.e. the cultural groups do also assimilate behaviorally.

What can also be seen in the present application, although the share of the minority group is only one fifth of the total population, the socialization investments of its parents succeed those of the majority's parents to a high enough extent for the cultural system to converge to a steady state culture that is located in the interior of the population–share weighted convex combination of the two initial cultures and the initial culture of the minority. The higher incentives of the parents of the minority group to invest into their socialization instruments stem from the more unfavorable cultural composition of the general socialization environment of their children (the unrelated adults), which makes it less suitable to use for substitution of own socialization investments. Finally, compared to the left case of Figure 1, the right case features a higher socialization value for both cultural groups. This leads into strictly larger socialization investments of the parents of both cultural groups for any given cultural distance (Proposition 5), so that the process of cultural and behavioral assimilation is accordingly slower.

5 Conclusions

This paper extended and generalized existing approaches in the cultural transmission literature to the formation of (continuous) cultural values (cultural preferences or traits) that we interpret as collectively defining the culture of a person. Under this property, we can define a collection of (adult) individuals that have a homogeneous culture as a *cultural group*. Within the endogenous cultural formation of preferences framework, that this paper has introduced as its first main part (which is introduced and summarized in section 1.3), there then arise the natural questions of qualitative properties of the behavioral and cultural assimilation processes between different cultural groups. To analyze these issues has been the task of the second main part of the present paper.

Thereby, the analysis is restricted to the case of two different cultural groups, where the homogeneous cultures of the adults of a group are specified to one collective endogenous cultural value and one exogenously fixed socialization value (i.e. the collective value captures all cultural information other than the socialization value). As far as the latter is concerned, we consider different classes of socialization values that determine the target for the future (demonstrated) culture of the children.

Since one of the main objects of the analysis is the endogenous evolution of the cultural distance between the endogenous cultural values of the two groups (which we define as the respective Euclidean distance), we introduce suitable conditions that guarantee that in any equilibrium of a given period, all parents of a cultural group take identical decisions with respect to the demonstrated culture and the socialization success rate. The conditions also ensure that all children of a cultural group do adopt the same culture and, thus, we obtain a continuous inter–generational representation of the two cultural groups.

The most important results that hold in a (weak) equilibrium of a period under a general class of socialization values are the following. In case of a strictly positive cultural distance between the two groups, parents always choose a behavioral dis–integration, complemented with a strictly positive socialization success rate. This comes in an attempt to countervail the 'negative' cultural influence that the social learning from the other group's adults has on the own child. Further, the relative positions of the two cultural groups in terms of a lower or higher endogenous cultural value are sustained over generations, i.e. the 'labels' of the two cultural groups are dynamically persistent.

As far as the comparative statics are concerned, we can show that if the two socialization instruments are not too strong substitutes in the parental optimization problem, then both the behavioral dis–integration, as well as the investments in the socialization success rate, are increasing in the socialization value. The same effect holds for (symmetric) increases in the cultural distance between the two groups. In both cases then, the cultural assimilation process between the two groups is slower (or an eventual cultural dissimilation process is faster). This theoretical result might deliver an explanation for the empirical observation of different paths of the integration and assimilation processes of cultural groups, since it both depends on the (perceived) cultural distance between the groups, as well as the strength of inter–generational considerations (the socialization value).

Notably, these theoretical results constitute empirically testable hypotheses. In this respect, the main explanatory variables for a cultural and behavioral assimilation process are the type and strength of the socialization value, as well as the (initial) cultural distance. As far as the first is concerned, good proxies to represent a general class of socialization values would e.g. be the 'importance of family', combined with the 'strength of behavioral norms' in a cultural group[38]. Since we constructed demonstrated cultures as the image of socio–economic choices, an eventual behavioral assimilation process between cultural groups is directly empirically observable. Given the theoretical results of the present paper (notably the characterizations of an equilibrium of a period and the comparative statics results), the respective data can further serve as a good proxy for the underlying cultural distances and, thus, for the cultural assimilation or dissimilation process — in case that explicit data on the cultures of groups are not or not to a satisfactory

[38]Arnett (1995) labels the two polar cases of the strength of the behavioral norms in the socialization process as 'broad and narrow socialization'. "Cultures characterized by broad socialization encourage individualism, independence, and self–expression [. . .] . In contrast, cultures characterized by narrow socialization hold obedience and conformity as the highest values and discourage deviation from cultural expectations [. . .] ." (p. 617)

extent available.

To analyze the endogenous evolution of the cultural distance in a dynamic setting, we applied a special class of socialization values, which is based on a parental desire for *inter–generational cultural closeness*. The characteristic of this class is that, under a certain parental myopia, the socialization target for the children's future culture coincides with the culture of the parents. Notably, this class also covers the concept of *imperfect empathy*, as used by Bisin and Verdier (2000, 2001) and throughout the literature on the economics of cultural transmission. We can show that under the inter–generational cultural closeness socialization values, in case of a strictly positive cultural distance, the cultures that the children of both cultural groups adopt lie in the interior of the interval that is constituted by the cultures of both groups. Further, in case of equal cultures, the parents do have no need to invest into their socialization instruments, so that such a homogeneous culture is a steady state. These results imply that if the cultural system starts out with initially different cultures, the cultural distance strictly declines over generations (cultural assimilation). Even, the two cultures converge to a homogeneous steady state, and the resulting steady state culture can be interpreted as a mixture of the two original cultures. This corresponds to the 'melting pot' theory of integration of cultural groups.

There are a number of directions, in which the present theoretical analysis could be extended. First, it might be interesting to identify whether there exist classes of socialization values under which non–decreasing cultural distances, and a cultural equilibrium with strictly positive cultural distance, are possible. Second, the behavior of the cultural system under different classes of socialization values should also be analyzed in a generalized culture framework, i.e. one with an n–dimensional set of cultural values, eventually combined with an endogenous treatment of the socialization value(s).

References

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Arnett, J. J. (1995). Broad and narrow socialization: The family in the context of a cultural theory. *Journal of Marriage and the Family*, 57:617–628.

Bandura, A. (1977). *Social Learning Theory.* General Learning Press, New York, NY.

Bandura, A. and Walters, R. H. (1963). *Social Learning and Personality Development.* Holt, Rinehart & Winston, New York.

Bisin, A., Patacchini, E., Verdier, T., and Zenou, Y. (2006). 'bend it like beckham': Identity, socialization and assimilation. CEPR Discussion Papers 5662, C.E.P.R. Discussion Papers.

Bisin, A., Patacchini, E., Verdier, T., and Zenou, Y. (2008). Are muslim immigrants different in terms of cultural integration? *Journal of the European Economic Association*, 6(2-3):445–456.

Bisin, A. and Topa, G. (2003). Empirical models of cultural transmission. *Journal of the European Economic Association*, 1(2-3):363–375.

Bisin, A. and Verdier, T. (1998). On the cultural transmission of preferences for social status. *Journal of Public Economics*, 70(1):75–97.

Bisin, A. and Verdier, T. (2000). A model of cultural transmission, voting and political ideology. *European Journal of Political Economy*, 16(1):5–29.

Bisin, A. and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97(2):298–319.

Bisin, A. and Verdier, T. (2008). Cultural transmission. In Durlauf, S. N. and Blume, L. E., editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, Basingstoke.

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

Brueckner, J. K. and Smirnov, O. (2007). Workings of the melting pot: Social networks and the evolution of population attributes. *Journal of Regional Science*, 47(2):209–228.

Brueckner, J. K. and Smirnov, O. (2008). Social networks and the convergence of population attributes: A generalization. *Journal of Regional Science*, 48(2):359–365.

Cavalli-Sforza, L. L. and Feldman, M. (1973). Cultural versus biological inheritance: phenotypic transmission from parents to children. *American Journal of Human Genetics*, 25:618–637.

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

Festinger, L. (1957). *A theory of cognitive dissonance*. Row, Peterson; Evanston/III.

Gordon, M. M. (1964). *Assimilation in American Life*. Oxford University Press, New York.

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

Harsanyi, J. C. (1982). Morality and the theory of rational behavior. In Sen, A. and Williams, B., editors, *Utilitarianism and beyond*, pages 39–62. Cambrigde: Cambridge University Press.

Hauk, E. and Sáez-Martí, M. (2002). On the cultural transmission of corruption. *Journal of Economic Theory*, 107(2):311–335.

Olcina, G. and Penarrubia, C. (2004). Hold up and intergenerational transmission of preferences. *Journal of Economic Behavior & Organization*, 54(1):111–132.

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.

Panebianco, F. (2009). 'driving while black': A theory for interethnic integration and evolution of prejudice. Working Papers 2009 10, University of Venice 'Ca' Foscari', Department of Economics.

Rotter, J. B. (1954). *Social learning and clinical psychology.* Prentice Hall, Englewood Cliffs/N.J.

Sáez-Martí, M. and Sjögren, A. (2007). Peers and culture. IEW - Working Papers iewwp349, Institute for Empirical Research in Economics - IEW.

Sen, A. (1977). Rational fools: A critique of the behavioral foundations of economic theory. *Philosophy and Public Affairs*, Summer:317–44.

Skinner, B. F. (1938). *The behavior of organisms.* Appleton, New York.

Skinner, B. F. (1953). *Science and human behavior.* Macmillan, New York.

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

A Appendix

A.1 Extensions

A.1.1 Properties of the Socialization Success Functions

Let us first discuss the specification of the determinants of the socialization successes in an top–down approach (concerning the level of detail of the representation). To the best of the author's assessment, the most general representation of the latter views the socialization successes as a function of the extensity and intensity with which a person socializes a child, $\sigma^n_{ac} := \sigma(\tau_{ac}, \iota^n_{ac})$, $\forall (a, c, n) \in \mathcal{A} \times \mathcal{C} \times \mathcal{N}$. The extensity of socialization is represented by $\tau_{ac} \in [0, 1]$ (without loss of generality, we normalize the length of the socialization period to one), which coincides with the length of the time that the person has spent with the child and is interpreted as a global determinant of the socialization successes of all cultural values, while as the socialization intensities, $\iota^n_{ac} \in \mathbb{R}_+$, are viewed as cultural value specific (this issue will be discussed in more detail below).

It is immediate to assume that if a child has no contact with a person, then it has no socialization influence, $\sigma(0, \cdot) = 0$. In line with the basic assumption in section 2, we also assume that the socialization success can not be negative[39], $\sigma(\cdot, \cdot) \in \mathbb{R}_+$. Furthermore, we assume that if a child merely observes a demonstrated cultural value (which we normalize to correspond to a socialization intensity level of zero), this also yields strictly positive social learning effects, i.e. $\sigma(\cdot, 0)$ is strictly increasing. Together with the assumption that the aggregate exposure of the child to socialization influences coincides with the length of the socialization period, $\forall c \in \mathcal{C}$, $\int_{a\in\mathcal{A}} \tau_{ac} da = 1$, it follows that the aggregated socialization success is strictly positive, $\int_{a\in\mathcal{A}} \sigma(\tau_{ac}, \iota^n_{ac}) da > 0$, as well as that the socialization success rates of all persons lie in the unit interval, $\hat{\sigma}^n_{ac} \in [0, 1]$, $\forall a \in \mathcal{A}$, such that $\int_{a\in\mathcal{A}} \hat{\sigma}^n_{ac} da = 1$ (which corresponds to the basic properties of the socialization successes that we introduced in section 2).

Let us now discuss general properties of the cultural value specific intensities of socialization, ι^n_{ac}. We assume that they are determined by two basic factors: The general devotion of a person to a child, $\delta_{ac} \in \mathbb{R}_+$, which we interpret as a global determinant of all socialization intensities (and it also determines the trust that a child has in a person); and the credibility, $\chi^n_{ac} \in \mathbb{R}_+$, that a child assigns to the claim that is implied by a chosen

[39] The assumption of positive socialization successes comes at the cost of a loss of generality. In principle, it is possible that adult agents do cause a 'socialization unsuccess' in the socialization process of a child, e.g. if a demonstrated cultural value is immoral or 'against the human nature' (which can thus also be realized by a child with yet undefined cultural values) — and hence their demonstrated cultural value can be interpreted as a role–model of 'how not to behave'. Nevertheless, we do for simplicity not integrate the related phenomena into the formal representation of the present paper.

demonstrated cultural value of a person that this choice is really superior to all other possible choices of demonstrated cultural values. Thus, $\iota^n_{ac} := \iota\left(\delta_{ac}, \chi^n_{ac}\right)$.

As far as the credibility levels are concerned, they themselves depend crucially on two factors. First, a cultural value specific socialization effort (e.g. the effort to convince a child about the superiority of adopting a certain realization of a cultural value compared to the other demonstrated cultural values that a child is confronted with in the socialization environment), $e^n_{ac} \in \mathbb{R}_+$. And second (as has already been noted in section 3), it is well known that individuals that seem to be more satisfied with their chosen life–style (demonstrated cultural values) have a more pronounced social learning impact in the children's socialization process. In light of the properties of the adults' preferences as introduced in the preceding subsection, the level of satisfaction crucially depends on the chosen demonstrated cultural values of the adults compared to their adopted cultural values. Thus, we finally obtain the cultural value specific credibilities as a function $\chi^n_{ac} := \chi\left(e^n_{ac}, \phi^n_a, \phi^{nd}_a\right)$.

Assumption 7 (Boundedness and Continuity)**.** $\forall (a, n) \in \mathcal{A} \times \mathcal{N}$

(a) the inputs to the socialization success functions, e^n_{ac} and δ_{ac}, are bounded, and

(b) the maps χ^n_{ac} and ι^n_{ac} are continuous, and $\sigma(\cdot, \cdot) = 0$ only at $\sigma(0, \cdot)$.

Under the latter assumption, the range of the map σ is also bounded. We finally use the properties of the socialization success functions (and their inputs) that we introduced in the present section to discuss in more detail the 'silent assumption' in section 3 that parents can choose all socialization success rates within the full unit interval, i.e. $\forall a \in \mathcal{A}$, $\hat{\Sigma}_a \in [0, 1]^n$. To do this, let Assumption 7 hold.

First, $\hat{\Sigma}_a = \mathbf{0}$ iff $\tau_{a\tilde{a}} =: \tau_a = 0$ since $\sigma(\cdot, \cdot) = 0$ only at $\sigma(0, \cdot)$ and $\int_{a' \in \mathcal{A}_a} \sigma\left(\tau_{a'\tilde{a}}, \cdot\right) da' < \infty$, $\forall \int_{a' \in \mathcal{A}_a} \tau_{a'\tilde{a}} = 1 - \tau_a$. This implies that parents can choose a zero interaction time with their children, i.e. effectively parents might 'give their children away' — which obviously constitutes a strong assumption.

Second, $\hat{\Sigma}_a = \mathbf{1}$ iff $\tau_a = 1$ since then $\sigma^n_{a'\tilde{a}} = 0$, $\forall a' \in \mathcal{A}_a$, and since $\int_{a' \in \mathcal{A}_a} \sigma\left(\tau_{a'\tilde{a}}, \cdot\right) da' > 0$, $\forall \int_{a' \in \mathcal{A}_a} \tau_{a'\tilde{a}} = 1 - \tau_a > 0$ and the map σ is bounded (so that also the parents can not achieve infinite socialization success). This implies that parents can choose to fully distract their children from the unrelated adults — which does obviously also constitute a strong assumption.

A.2 Proofs

A.2.1 Proof of Proposition 1

In this proof, the following further notation will be useful for brevity: $\hat{\Sigma}_{-a} := \left\{\hat{\Sigma}_{a'}\right\}_{a' \in \mathcal{A}_a}$, $\Omega^d_{-a} := \times_{a' \in \mathcal{A}_a} \Omega^d$, $[0, 1]^n_{-a} := \times_{a' \in \mathcal{A}_a} [0, 1]^n$. Further, note that under Assumption 3, the set $\Omega^d_{\mathcal{A}} \times [0,1]^n_{\mathcal{A}}$ is compact and convex since it is the Cartesian product of compact and convex sets (if the set of adults is a continuum, thus infinite, the product of the choice sets is endowed with the product topology).

(a)

Denote $\max_{\Phi^d_a \in \Omega^d} \mathcal{M}\left(\Phi^d_a, \hat{\Sigma}_a, \Phi^d_{-a}\right)$ as $\mathcal{M}_{a1}$ and $\max_{\hat{\Sigma}_a \in [0,1]^n} \mathcal{M}\left(\Phi^d_a, \hat{\Sigma}_a, \Phi^d_{-a}\right)$ as $\mathcal{M}_{a2}$. Next, define the 'best–reply' correspondence of problem $\mathcal{M}_{a1}$ as $B_a : [0,1]^n \times \Omega^d_{-a} \mapsto \Omega^d$ and the solution correspondence to problem $\mathcal{M}_{a2}$ as $S_a : \Omega^d_{\mathcal{A}} \mapsto [0,1]^n$. Denote the corresponding product correspondences as $B_{\mathcal{A}} = \times_{a \in \mathcal{A}} B_a : [0,1]^n_{\mathcal{A}} \times \Omega^d_{\mathcal{A}} \mapsto \Omega^d_{\mathcal{A}}$, respectively $S_{\mathcal{A}} = \times_{a \in \mathcal{A}} S_a : \Omega^d_{\mathcal{A}} \mapsto [0,1]^n_{\mathcal{A}}$. Finally, consider the composition of the product correspondences as $S_{\mathcal{A}} \circ B_{\mathcal{A}} : [0,1]^n_{\mathcal{A}} \times \Omega^d_{\mathcal{A}} \mapsto \Omega^d_{\mathcal{A}} \mapsto [0,1]^n_{\mathcal{A}}$, which can equivalently be written as $S_{\mathcal{A}} \circ B_{\mathcal{A}} : [0,1]^n_{\mathcal{A}} \times \Omega^d_{\mathcal{A}} \mapsto [0,1]^n_{\mathcal{A}} \times \Omega^d_{\mathcal{A}}$.

Since $\mathcal{M}_{a1}$ and $\mathcal{M}_{a2}$ are continuous and Ω^d and $[0,1]^n$ are 'constant' (and therefore continuous), it follows by Berge's Theorem of the Maximum that B_a and S_a are non–empty, compact–valued and upper hemicontinuous. Since both optimization problems are also strictly concave, the solution correspondences are even single–valued, and thus, B_a and S_a are continuous functions. These properties are inherited by the product correspondences $B_{\mathcal{A}}$ and $S_{\mathcal{A}}$. Finally, since $S_{\mathcal{A}} \circ B_{\mathcal{A}}$ is the composition of continuous functions, it is also a continuous function.

Now, since $S_{\mathcal{A}} \circ B_{\mathcal{A}} : [0,1]^n_{\mathcal{A}} \times \Omega^d_{\mathcal{A}} \mapsto [0,1]^n_{\mathcal{A}} \times \Omega^d_{\mathcal{A}}$ is a continuous function from a non–empty, convex and compact set into itself, it follows by Brouwer's Fix Point Theorem, that $\exists \left(\hat{\Sigma}^*_{\mathcal{A}}, \Phi^{d^*}_{\mathcal{A}}\right) \in [0,1]^n_{\mathcal{A}} \times \Omega^d_{\mathcal{A}}$ such that $\left(\hat{\Sigma}^*_{\mathcal{A}}, \Phi^{d^*}_{\mathcal{A}}\right) = S_{\mathcal{A}} \circ B_{\mathcal{A}} \left(\hat{\Sigma}^*_{\mathcal{A}}, \Phi^{d^*}_{\mathcal{A}}\right)$. □

Remark: Concavity or quasi–concavity of the problems $\mathcal{M}_{a1}$ and $\mathcal{M}_{a2}$ would not suffice to guarantee the existence of a fixed point of the mapping $SB_{\mathcal{A}}$. This follows since although the mappings B_a and S_a would be non–empty, compact–valued and UHC by Berge's Theorem of the Maximum, and furthermore convex–valued given concavity or quasi–concavity, the properties of which would be inherited by $B_{\mathcal{A}}$ and $S_{\mathcal{A}}$, the composite correspondence $S_{\mathcal{A}} \circ B_{\mathcal{A}}$ would neither be guaranteed to be compact–valued nor convex–valued (which would be needed to apply Kakutani's Fix Point Theorem).

(b)

Denote the best–reply correspondence of optimization problem (3) as $Y_a : \Omega^d_{-a} \times [0,1]^n_{-a} \mapsto \Omega^d \times [0,1]^n$ (optimization problem (3) is independent, and thus 'constant' and continuous with respect to the non–strategic choices of the other adults, $\hat{\Sigma}_{-a} \in [0,1]^n_{-a}$). Since the maximization problems (3) are continuous, it follows by Berge's Theorem of the Maximum that the best–

reply correspondences are non–empty, compact–valued and upper hemicontinuous, and further convex–valued (since the maximization problems (3) are quasi–concave). These properties are inherited by the product correspondence $Y_{\mathcal{A}} = \times_{a\in\mathcal{A}} Y_a : \Omega^d_{\mathcal{A}} \times [0,1]^n_{\mathcal{A}} \mapsto \Omega^d_{\mathcal{A}} \times [0,1]^n_{\mathcal{A}}$.

Since $Y_{\mathcal{A}}$ is a non–empty, compact– and convex–valued UHC correspondence from a non–empty, compact convex set into itself, it follows from Kakutani's Fix Point Theorem that $\exists \left(\Phi^{d^{**}}_{\mathcal{A}}, \hat{\Sigma}^{**}_{\mathcal{A}}\right) \in \Omega^d_{\mathcal{A}} \times [0,1]^n_{\mathcal{A}}$ such that $\left(\Phi^{d^{**}}_{\mathcal{A}}, \hat{\Sigma}^{**}_{\mathcal{A}}\right) \in Y_{\mathcal{A}} \left(\Phi^{d^{**}}_{\mathcal{A}}, \hat{\Sigma}^{**}_{\mathcal{A}}\right)$. □

A.2.2 Proof of Proposition 2

Under Assumption 1—4, all statements and results of the Proof of Proposition 1 (a) stay valid. Let us restrict the attention to the 'homogeneous subset' of $\Omega^d_{\mathcal{A}} \times [0,1]^n_{\mathcal{A}}$, i.e. the one where it holds that $\forall G \in \{L, H\}$, $\forall g \in \mathcal{G}$, $\left(\Phi^d_g, \hat{\Sigma}_g\right) = \left(\Phi^d_G, \hat{\Sigma}_G\right)$, thus we consider the set $diag\left(\Omega^d_{\mathcal{L}} \times [0,1]^n_{\mathcal{L}}\right) \times diag\left(\Omega^d_{\mathcal{H}} \times [0,1]^n_{\mathcal{H}}\right)$ (where the notation here is analogous to previous notation). Since this is the Cartesian product of the diagonals of non–empty, convex and compact sets, it inherits these properties.

On this 'homogeneous subset', it follows that under condition (5), the relative–success–weighted demonstrated cultures of the unrelated socialization environment, equations (1), are equal for all children (and thus enter equally in the parental socialization techniques (2)). Since also all utility and cost functions are equal for the adults of a cultural group, it further holds that the optimization problems $\mathcal{M}_{g1}$ and $\mathcal{M}_{g2}$ are identical for all $g \in \mathcal{G}$, $G \in \{L, H\}$, and thus, they have identical sets of optimizers. Since the optimization problems are furthermore strictly concave, and thus, the set of optimizers single–valued, it is also guaranteed that all adults of group 'select' identical elements (thus the 'homogeneous subset' maps into the 'homogeneous subset').

In analogy to the Proof of Proposition 1 (a) we thus have a continuous function from a non–empty, compact and convex set into itself

$$diag\left(\Omega^d_{\mathcal{L}} \times [0,1]^n_{\mathcal{L}}\right) \times diag\left(\Omega^d_{\mathcal{H}} \times [0,1]^n_{\mathcal{H}}\right) \mapsto diag\left(\Omega^d_{\mathcal{L}} \times [0,1]^n_{\mathcal{L}}\right) \times diag\left(\Omega^d_{\mathcal{H}} \times [0,1]^n_{\mathcal{H}}\right).$$

By Brouwer's Fixed Point Theorem, a fixed point of this mapping, which constitutes a HWEP, exists. □

A.2.3 Proof of Proposition 3

The optimization problem of any adult $g \in \mathcal{G}$ of any cultural group $G \in \{L, H\}$ is equation (4), subject to equation (2). The corresponding Lagrangeans are

$$\begin{aligned} \mathcal{L}_g\left(\phi^d_g, \hat{\sigma}_g\right) = {} & u^{\phi_G}\left(\phi^d_g\right) + s^{\phi^s_G}\left(\phi_{\tilde{g}}, \hat{\phi}_G\right) - c\left(\hat{\sigma}_g, u^{\phi_G}(\phi^d_g)\right) \\ & + \lambda_0(\phi^d_g - \underline{\phi}) + \lambda_1(\overline{\phi} - \phi^d_g) + \mu_0 \hat{\sigma}_g + \mu_1(1 - \hat{\sigma}_g) \end{aligned} \quad \text{(A.1)}$$

The first order conditions (which do necessarily have to hold both in a HWEP and a HEP), $\frac{\partial \mathcal{L}_g(\phi_g^d,\hat{\sigma}_g)}{\partial \phi_g^d} = 0$ and $\frac{\partial \mathcal{L}_g(\phi_g^d,\hat{\sigma}_g)}{\partial \hat{\sigma}_g} = 0$, evaluated under condition (5) and the homogeneity requirement of a HWEP, under which $\forall G \in \{L, H\}$, $\forall g \in \mathcal{G}$, $\phi_g^d = \phi_G^d$ and $\hat{\sigma}_g = \hat{\sigma}_G$, and then $\phi_{\tilde{g}} = \phi_{\tilde{G}} = \phi_G^d + (1-\hat{\sigma}_G)\left(q_G\phi_G^d + (1-q_G)\phi_{-G}^d\right)$, are

$$\frac{\partial\, u^{\phi_G}\left(\phi_G^d\right)}{\partial\, \phi_G^d}\left(1 - \frac{\partial\, c\left(\hat{\sigma}_G, u^{\phi_G}\left(\phi_G^d\right)\right)}{\partial\, u^{\phi_G}\left(\phi_G^d\right)}\right) + \frac{\partial\, \mathrm{s}^{\phi^{\mathrm{s}}_{\mathrm{G}}}\left(\phi_{\tilde{\mathrm{G}}}, \hat{\phi}_{\mathrm{G}}\right)}{\partial\, \phi_{\tilde{\mathrm{G}}}}\hat{\sigma}_G + \lambda_0 - \lambda_1 = 0 \tag{A.2}$$

$$\frac{\partial\, \mathrm{s}^{\phi^{\mathrm{s}}_{\mathrm{G}}}\left(\phi_{\tilde{\mathrm{G}}}, \hat{\phi}_{\mathrm{G}}\right)}{\partial\, \phi_{\tilde{\mathrm{G}}}}(1-q_G)(\phi_G^d - \phi_{-G}^d) - \frac{\partial\, c\left(\hat{\sigma}_G, u^{\phi_G}\left(\phi_G^d\right)\right)}{\partial\, \hat{\sigma}_G} + \mu_0 - \mu_1 = 0 \tag{A.3}$$

For the subsequent analysis it will be useful to remember that by Assumptions 1 (b) and 5 (a), $\frac{\partial\, u^{\phi_G}(\phi_G^d)}{\partial\, \phi_G^d}$, respectively $\frac{\partial\, \mathrm{s}^{\phi^{\mathrm{s}}_{\mathrm{G}}}(\phi_{\tilde{\mathrm{G}}},\hat{\phi}_{\mathrm{G}})}{\partial\, \phi_{\tilde{\mathrm{G}}}}$, are positive/negative/zero whenever ϕ_G^d, respectively $\phi_{\tilde{G}}$, are smaller/larger/equal ϕ_G, respectively $\hat{\phi}_G$.

1. Case $\phi_H > \phi_L$

(a) $\hat{\sigma}_G^* > 0$:

Assume $\hat{\sigma}_G = 0$. Then, FOC (A.3) is $\frac{\partial\, \mathrm{s}^{\phi^{\mathrm{s}}_{\mathrm{G}}}(\phi_{\tilde{\mathrm{G}}},\hat{\phi}_{\mathrm{G}})}{\partial\, \phi_{\tilde{\mathrm{G}}}}(1-q_G)\left(\phi^{\mathrm{d}}_{\mathrm{G}}\text{-}\phi^{\mathrm{d}}_{\text{-G}}\right) + \mu_0 = 0$. Note that if $\hat{\sigma}_G = 0$, parents will optimally set $\phi_G^d = \phi_G$, since their demonstrated culture is fully ineffective in the socialization process. In case $\mu_0 = 0$, we hence require $\phi_{\tilde{G}} = \hat{\phi}_G$, which implies $\phi_{-G}^d = \hat{\phi}_G$, and/or $\phi_{-G}^d = \phi_G^d = \phi_G$. But for parents $-G$ both demonstrated cultures would mean $\phi_{-G}^d > (<)\phi_{-G}$ and $\tilde{\phi}_{-G} > (<)\hat{\phi}_{-G}$, which would yield, $\forall -g \in -\mathcal{G}$, $\frac{\partial\, \mathcal{L}_{-g}(\phi_{-G}^d,\hat{\sigma}_{-G})}{\partial\, \phi_{-g}^d} < (>)0$. The case $\mu_0 > 0$ would require $\frac{\partial\, \mathrm{s}^{\phi^{\mathrm{s}}_{\mathrm{G}}}(\phi_{\tilde{\mathrm{G}}},\hat{\phi}_{\mathrm{G}})}{\partial\, \phi_{\tilde{\mathrm{G}}}}(1-q_G)\left(\phi^{\mathrm{d}}_{\mathrm{G}}\text{-}\phi^{\mathrm{d}}_{\text{-G}}\right) < 0$. For this to hold, the pairs $\phi_{\tilde{G}} > (<)\hat{\phi}_G$ and $\phi_G^d > (<)\phi_{-G}^d$ are needed. Consider the type H parents. $\phi_L^d = \phi_{\tilde{H}} > \hat{\phi}_H \geq \phi_H = \phi_H^d$ does not fulfill the desired relationships. The same is true for $\phi_L^d = \phi_{\tilde{H}} < \phi_H = \phi_H^d$. If $\phi_L^d = \phi_{\tilde{H}} \in [\phi_H, \hat{\phi}_H]$, then $\phi_L^d > \phi_L$ and $\phi_{\tilde{L}} > \phi_L$ which implies, $\forall l \in \mathcal{L}$, $\frac{\partial\, \mathcal{L}_l(\phi_L^d,\hat{\sigma}_L)}{\partial\, \phi_l^d} < 0$. Similar, for the type L parents, $\phi_H^d = \phi_{\tilde{L}} < \hat{\phi}_L \leq \phi_L = \phi_L^d$ and $\phi_H^d = \phi_{\tilde{L}} > \phi_L = \phi_L^d$ do not satisfy the desired relationships. If $\phi_H^d = \phi_{\tilde{L}} \in [\hat{\phi}_L, \phi_L]$, then $\phi_H^d < \phi_H$ and $\phi_{\tilde{H}} > \phi_H$ which implies, $\forall h \in \mathcal{H}$, $\frac{\partial\, \mathcal{L}_h(\phi_H^d,\hat{\sigma}_H)}{\partial\, \phi_h^d} > 0$. We can conclude that $\hat{\sigma}_G^* \in (0, 1]$ in a HWEP. □

(b)+(c) $\phi_H^{d^*} \geq \phi_H > \phi_L \geq \phi_L^{d^*}$, $\hat{\phi}_H > \phi_{\tilde{H}}^* > \phi_{\tilde{L}}^* > \hat{\phi}_L$:

Case $\phi_G \in \left(\underline{\phi}, \overline{\phi}\right)$, $\forall G \in \{L, H\}$: Assume $\phi_G^d = \phi_G$. In this case, FOC (A.2) reduces to $\frac{\partial s^{\phi^S_G}(\phi_{\tilde{G}}, \hat{\phi}_G)}{\partial \phi_{\tilde{G}}} \hat{\sigma}_G = 0$, and we require $\phi_{\tilde{G}} = \hat{\phi}_G$ (since $\hat{\sigma}_G > 0$ in a HWEP). But then $\frac{\partial \mathcal{L}_g(\phi_G^d, \hat{\sigma}_G)}{\partial \hat{\sigma}_G} < 0$. Assume $\phi_G^d > (<) \phi_G$. Then, for FOC (A.2) to hold, $\frac{\partial s^{\phi^S_G}(\phi_{\tilde{G}}, \hat{\phi}_G)}{\partial \phi_{\tilde{G}}} > (<) 0$ is required, which implies $\phi_{\tilde{G}} < (>) \hat{\phi}_G$. Hence, only the combination of the relations $\phi_G^d > (<) \phi_G$ and $\phi_{\tilde{G}} < (>) \hat{\phi}_G$ can realize in a HWEP. We need to show that exactly $\phi_H^d > \phi_H > \phi_L > \phi_L^d$ and $\hat{\phi}_H > \phi_{\tilde{H}} > \phi_{\tilde{L}} > \hat{\phi}_L$ holds. Suppose that this does not hold. Then $\phi_L^d > \phi_L$, $\phi_H^d < \phi_H$, $\phi_{\tilde{L}} < \hat{\phi}_L$ and $\phi_{\tilde{H}} > \hat{\phi}_H$. For this to be possible, it is needed that $\phi_L^d \geq \phi_{\tilde{H}} > \phi_{\tilde{L}} \geq \phi_H^d$ since both future cultures are a convex combination of the demonstrated cultures of parents of both cultural groups. But since $(1 - q_H)\phi_L^d + q_H \phi_H^d := \phi_A^d \in (\phi_H^d, \phi_L^d)$ and $\hat{\sigma}_G > 0$, $\forall G \in \{L, H\}$, it would also have to hold that $\hat{\sigma}_L \phi_L^d + (1 - \hat{\sigma}_L)\phi_A^d = \phi_{\tilde{L}} > \phi_{\tilde{H}} = \hat{\sigma}_H \phi_H^d + (1 - \hat{\sigma}_H)\phi_A^d$, which yields a contradiction.

Case $\phi_L = \underline{\phi}$ and/or $\phi_H = \overline{\phi}$: Suppose $\phi_L^d > \phi_L = \underline{\phi}$ $\left(\phi_H^d < \phi_H = \overline{\phi}\right)$. Then, for FOC (A.2) to hold, we require $\frac{\partial s^{\phi^S_G}(\phi_{\tilde{G}}, \hat{\phi}_G)}{\partial \phi_{\tilde{G}}} \hat{\sigma}_G > (<) 0$, $G = L(H)$, which implies $\phi_{\tilde{L}} < \hat{\phi}_L = \phi_L = \underline{\phi}$ ($\phi_{\tilde{H}} > \hat{\phi}_H = \phi_H = \overline{\phi}$), which is not possible.

It follows that in a HWEP $\phi_H^{d*} \geq \phi_H > \phi_L \geq \phi_L^{d*}$ (with equalities iff $\phi_H = \overline{\phi}$ or $\phi_L = \underline{\phi}$) and $\hat{\phi}_H > \phi_{\tilde{H}}^* > \phi_{\tilde{L}}^* > \hat{\phi}_L$. □

2. Case $\phi_H = \phi_L$

(b) $\phi_H^{d*} \geq \phi_H = \phi_L \geq \phi_L^{d*}$:
Let us assume that $\phi_H^d < \phi_H$. Then, for FOC (A.2) to be fulfilled for the H parents, it must hold that $\phi_{\tilde{H}} > \hat{\phi}_H$. This necessarily requires $\phi_L^d > \hat{\phi}_H \geq \phi_H = \phi_L$, and then for the FOC (A.2) of the L parents to hold, $\phi_{\tilde{L}} < \hat{\phi}_L \leq \hat{\phi}_H$ is needed. Hence, the pairs $\phi_{\tilde{H}} > \phi_{\tilde{L}}$ and $\phi_H^d < \phi_L^d$ are required in a hypothetical optimum. But this yields a contradiction, since the final cultures of the children of both groups are convex combinations of the representative demonstrated culture of the public social space, and the demonstrated cultures of the parents (and hence the order of the demonstrated cultures of the parents of the two groups must be weakly preserved for the cultures of the children). It is immediate, that the same contradiction would follow, if we initially assumed that $\phi_L^d > \phi_L$.

(c) $\hat{\phi}_H \geq \phi_{\tilde{H}}^* \geq \phi_{\tilde{L}}^* \geq \hat{\phi}_L$:
If $\phi_H^{d*} = \phi_L^{d*}$, then the result holds trivially. In the case $\phi_H^{d*} > \phi_L^{d*}$ for FOC (A.3) to hold, it follows that $sign(\phi_{\tilde{G}}^* - \hat{\phi}_G) = -sign(\phi_G^{d*} - \phi_G)$, $\forall G \in \{L, H\}$. □

A.2.4 Proof of Proposition 4

The structure of this proof is analogous to the one in Appendix A.2.2. First note that under Assumptions 4 and 6 (a), the optimization problem (4) is concave (and thus quasi–concave and continuous) so that all results of the Proof of Proposition A.2.1 (b) stay valid. Now, consider any point on the 'homogeneous subset' $diag\left(\Omega^d_{\mathcal{L}} \times [0,1]^n_{\mathcal{L}}\right) \times diag\left(\Omega^d_{\mathcal{H}} \times [0,1]^n_{\mathcal{H}}\right)$. Under condition (5), the best–reply correspondences to this point are identical for all adult members of a group. Furthermore, by the results of Proof of Proposition A.2.1 (b), the set of 'homogeneous best replies' (i.e. the intersection of the product of the individual best replies, $Y_{\mathcal{A}}$, with the 'homogeneous subset') is non–empty, convex– and compact–valued and UHC.

We thus can construct a non–empty, convex– and compact–valued, UHC correspondence from a compact and convex set into itself

$$diag\left(\Omega^d_{\mathcal{L}} \times [0,1]^n_{\mathcal{L}}\right) \times diag\left(\Omega^d_{\mathcal{H}} \times [0,1]^n_{\mathcal{H}}\right) \mapsto diag\left(\Omega^d_{\mathcal{L}} \times [0,1]^n_{\mathcal{L}}\right) \times diag\left(\Omega^d_{\mathcal{H}} \times [0,1]^n_{\mathcal{H}}\right)$$

and, thus, by Kakutani's Fixed Point Theorem, a fix point of this mapping, which constitutes a HEP, exists. □

A.2.5 Proof of Proposition 5

We assume subsequently that at the original solutions to the optimization problems, the constraints of the parental optimization problems (4) have not been binding, i.e. $\overline{\phi} > \phi_H^{d^{**}} > \phi_L^{d^{**}} > \underline{\phi}$ and $\hat{\sigma}_G^{**} \in (0,1)\ \forall G \in \{L,H\}$. At any HEP, the following signs for the second partial derivatives of the Lagrangean (A.1) with respect to the endogenous parental choice variables hold[40] (subsequently, the entries in the sign–matrices will always correspond to $G = L/H$)

$$\text{sign} \begin{pmatrix} \frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial {\phi^{\mathrm{d}}_{\mathrm{G}}}^2} & \frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}}\, \partial \hat{\sigma}_{\mathrm{G}}} & \frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}}\, \partial \phi^{\mathrm{d}}_{\text{-G}}} & \frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}}\, \partial \hat{\sigma}_{\text{-G}}} \\ \frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}}\, \partial \phi^{\mathrm{d}}_{\mathrm{G}}} & \frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial {\hat{\sigma}_{\mathrm{G}}}^2} & \frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}}\, \partial \phi^{\mathrm{d}}_{\text{-G}}} & \frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}}\, \partial \hat{\sigma}_{\text{-G}}} \end{pmatrix} = \begin{pmatrix} -/- & \sim/\sim & -/- & 0/0 \\ \sim/\sim & -/- & +/- & 0/0 \end{pmatrix} \tag{A.4}$$

where

$$\frac{\partial^2 \mathcal{L}^{**}_G}{\partial \phi^d_G\, \partial \hat{\sigma}_G} = \frac{\partial^2 s^{\phi^s_G}\left(\phi^{**}_{\tilde{G}}, \hat{\phi}_G\right)}{\partial {\phi_{\tilde{G}}}^2} \hat{\sigma}^{**}_G (1-q_G)(\phi^{d^{**}}_G - \phi^{d^{**}}_{-G}) + \frac{\partial s^{\phi^s_G}\left(\phi^{**}_{\tilde{G}}, \hat{\phi}_G\right)}{\partial \phi_{\tilde{G}}} - \frac{\partial^2 c\left(\hat{\sigma}^{**}_G, u^{\phi_G}\left(\phi^{d^{**}}_G\right)\right)}{\partial \hat{\sigma}_G\, \partial u^{\phi_G}\left(\phi^d_G\right)}$$

[40]For saving space, we will throughout this proof denote $\mathcal{L}^{**}_G := \mathcal{L}_g\left(\phi^{d^{**}}_G, \hat{\sigma}^{**}_G\right)$ (the Lagrangeans are equal in any HEP for all members of any cultural group).

the signs of which are in general ambiguous. The signs for the second partial derivatives of the Lagrangean (A.1) with respect to the state variables and parameters of the model are

$$
\text{sign}\begin{pmatrix}
\frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \phi_{\mathrm{G}}^{\mathrm{d}}\,\partial \phi_{\mathrm{G}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \phi_{\mathrm{G}}^{\mathrm{d}}\,\partial \phi_{\text{-}\mathrm{G}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \phi_{\mathrm{G}}^{\mathrm{d}}\,\partial \phi_{\mathrm{G}}^{\mathrm{s}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \phi_{\mathrm{G}}^{\mathrm{d}}\,\partial \phi_{\text{-}\mathrm{G}}^{\mathrm{s}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \phi_{\mathrm{G}}^{\mathrm{d}}\,\partial \hat{\phi}_{\mathrm{G}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \phi_{\mathrm{G}}^{\mathrm{d}}\,\partial \hat{\phi}_{\text{-}\mathrm{G}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \phi_{\mathrm{G}}^{\mathrm{d}}\,\partial \mathrm{q}_{\mathrm{G}}} \\
\frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \hat{\sigma}_{\mathrm{G}}\,\partial \phi_{\mathrm{G}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \hat{\sigma}_{\mathrm{G}}\,\partial \phi_{\text{-}\mathrm{G}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \hat{\sigma}_{\mathrm{G}}\,\partial \phi_{\mathrm{G}}^{\mathrm{s}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \hat{\sigma}_{\mathrm{G}}\,\partial \phi_{\text{-}\mathrm{G}}^{\mathrm{s}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \hat{\sigma}_{\mathrm{G}}\,\partial \hat{\phi}_{\mathrm{G}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \hat{\sigma}_{\mathrm{G}}\,\partial \hat{\phi}_{\text{-}\mathrm{G}}} & \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \hat{\sigma}_{\mathrm{G}}\,\partial \mathrm{q}_{\mathrm{G}}}
\end{pmatrix}
=
\begin{pmatrix}
+/+ & 0/0 & -/+ & 0/0 & +/+ & 0/0 & +/- \\
-/+ & 0/0 & +/+ & 0/0 & -/+ & 0/0 & -/-
\end{pmatrix}
\tag{A.5}
$$

By the Implicit Function Theorem[41], we obtain the following comparative statics effects of the state variables and parameters of the model in a HEP[42]

$$
\text{sign}\left(\frac{\partial \phi_G^{d^{**}}}{\partial \phi_G}\right) =
$$

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

$$
\text{sign}\left(\frac{\partial \sigma_G^{**}}{\partial \phi_G}\right) = \tag{A.6}
$$

$$
\begin{aligned}
\text{sign}\Bigg(&\left(\frac{\partial^2 \mathcal{L}_{\text{-}\mathrm{G}}^{**}}{\partial {\phi_{\text{-}\mathrm{G}}^{\mathrm{d}}}^2}\,\frac{\partial^2 \mathcal{L}_{\text{-}\mathrm{G}}^{**}}{\partial {\hat{\sigma}_{\text{-}\mathrm{G}}}^2} - \left(\frac{\partial^2 \mathcal{L}_{\text{-}\mathrm{G}}^{**}}{\partial \phi_{\text{-}\mathrm{G}}^{\mathrm{d}}\,\partial \hat{\sigma}_{\text{-}\mathrm{G}}}\right)^2\right)\left(\frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \phi_{\mathrm{G}}^{\mathrm{d}}\,\partial \phi_{\mathrm{G}}}\,\frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \phi_{\mathrm{G}}^{\mathrm{d}}\,\partial \hat{\sigma}_{\mathrm{G}}} - \frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial {\phi_{\mathrm{G}}^{\mathrm{d}}}^2}\,\frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \hat{\sigma}_{\mathrm{G}}\,\partial \phi_{\mathrm{G}}}\right) + \\
&\left(\frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial {\phi_{\mathrm{G}}^{\mathrm{d}}}^2}\,\frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial {\hat{\sigma}_{\mathrm{G}}}^2} - \left(\frac{\partial^2 \mathcal{L}_{\mathrm{G}}^{**}}{\partial \phi_{\mathrm{G}}^{\mathrm{d}}\,\partial \hat{\sigma}_{\mathrm{G}}}\right)^2\right)\left(\frac{\partial^2 \mathcal{L}_{\text{-}\mathrm{G}}^{**}}{\partial \phi_{\text{-}\mathrm{G}}^{\mathrm{d}}\,\partial \hat{\sigma}_{\text{-}\mathrm{G}}}\,\frac{\partial^2 \mathcal{L}_{\text{-}\mathrm{G}}^{**}}{\partial \hat{\sigma}_{\text{-}\mathrm{G}}\,\partial \phi_{\mathrm{G}}^{\mathrm{d}}} - \frac{\partial^2 \mathcal{L}_{\text{-}\mathrm{G}}^{**}}{\partial \phi_{\text{-}\mathrm{G}}^{\mathrm{d}}\,\partial \phi_{\mathrm{G}}^{\mathrm{d}}}\,\frac{\partial^2 \mathcal{L}_{\text{-}\mathrm{G}}^{**}}{\partial {\hat{\sigma}_{\text{-}\mathrm{G}}}^2}\right)\Bigg)
\end{aligned}
$$

---

[41]By Assumption 4 and 6 (a) together with the condition of Proposition 5, the parental optimization problems are strictly concave at the HEP under consideration, which implies that the determinant of the Hessian matrix is strictly positive and all conditions for the Implicit Function Theorem are fulfilled.

[42]Note here for completeness that from the analysis in the 'expanded' form of the second partial derivatives below, it follows that in any HEP the following results must hold without imposing the additional restriction of Proposition 5 $\begin{pmatrix} \frac{\partial \phi_G^{d^{**}}}{\partial \hat{\phi}_G} \\ \frac{\partial \hat{\sigma}_G^{*}}{\partial \hat{\phi}_G} \end{pmatrix} = -\frac{\frac{\partial^2 s^{\phi_G^s}\left(\phi_{\tilde{G}}^{**},\hat{\phi}_G\right)}{\partial {\phi_{\tilde{G}}}^2}}{\frac{\partial^2 s^{\phi_G^s}\left(\phi_{\tilde{G}}^{**},\hat{\phi}_G\right)}{\partial \phi_{\tilde{G}}\,\partial \phi_G^s}} \begin{pmatrix} \frac{\partial \phi_G^{d^{**}}}{\partial \phi_G^s} \\ \frac{\partial \hat{\sigma}_G^{**}}{\partial \phi_G^s} \end{pmatrix}$ and $\begin{pmatrix} \frac{\partial \phi_G^{d^{**}}}{\partial \hat{\phi}_{-G}} \\ \frac{\partial \hat{\sigma}_G^{**}}{\partial \hat{\phi}_{-G}} \end{pmatrix} = -\frac{\frac{\partial^2 s^{\phi_G^s}\left(\phi_{\tilde{G}}^{**},\hat{\phi}_G\right)}{\partial {\phi_{\tilde{G}}}^2}}{\frac{\partial^2 s^{\phi_G^s}\left(\phi_{\tilde{G}}^{**},\hat{\phi}_G\right)}{\partial \tilde{\phi}_{-G}\,\partial \phi_{-G}^s}} \begin{pmatrix} \frac{\partial \phi_G^{d^{**}}}{\partial \phi_{-G}^s} \\ \frac{\partial \hat{\sigma}_G^{**}}{\partial \phi_{-G}^s} \end{pmatrix}$ and thus for $G = L/H$, $\text{sign}\begin{pmatrix} \frac{\partial \phi_G^{d^{**}}}{\partial \hat{\phi}_G} \\ \frac{\partial \hat{\sigma}_G^{**}}{\partial \hat{\phi}_G} \end{pmatrix} = -/+\ \text{sign}\begin{pmatrix} \frac{\partial \phi_G^{d^{**}}}{\partial \phi_G^s} \\ \frac{\partial \hat{\sigma}_G^{*}}{\partial \phi_G^s} \end{pmatrix}$ and $\text{sign}\begin{pmatrix} \frac{\partial \phi_G^{d^{**}}}{\partial \hat{\phi}_{-G}} \\ \frac{\partial \hat{\sigma}_G^{**}}{\partial \hat{\phi}_{-G}} \end{pmatrix} = +/-\ \text{sign}\begin{pmatrix} \frac{\partial \phi_G^{d^{**}}}{\partial \phi_{-G}^s} \\ \frac{\partial \hat{\sigma}_G^{**}}{\partial \phi_{-G}^s} \end{pmatrix}$.

$$
\operatorname{sign}\left(\frac{\partial\,\phi_G^{d^{**}}}{\partial\,\phi_{-G}}\right) =
$$

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

$$
\operatorname{sign}\left(\frac{\partial\,\sigma_G^{**}}{\partial\,\phi_{-G}}\right) =
$$

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

$$
\operatorname{sign}\left(\frac{\partial\,\phi_G^{d^{**}}}{\partial\,\phi_G^{s}}\right) =
$$

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

$$
\operatorname{sign}\left(\frac{\partial\,\sigma_G^{**}}{\partial\,\phi_G^{s}}\right) = \tag{A.7}
$$

$$
\begin{aligned}
\operatorname{sign}\Bigg(&\left(\left(\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}\right)^2-\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}{}^2}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}{}^2}\right)\left(\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}{}^2}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\phi^{\text{s}}_{\text{G}}}-\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat{\sigma}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\phi^{\text{s}}_{\text{G}}}\right)+\\
&\left(\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\phi^{\text{d}}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}{}^2}-\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,\hat{\sigma}_{\text{-G}}\,\partial\,\phi^{\text{d}}_{\text{G}}}\right)\left(\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\phi^{\text{s}}_{\text{G}}}-\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\phi^{\text{s}}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\phi^{\text{d}}_{\text{-G}}}\right)\Bigg)
\end{aligned}
$$

$$
\operatorname{sign}\left(\frac{\partial\,\phi_G^{d^{**}}}{\partial\,\phi^{s}_{-G}}\right) =
$$

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

$$
\operatorname{sign}\left(\frac{\partial\,\sigma_G^{**}}{\partial\,\phi^{s}_{-G}}\right) =
$$

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

$$
\operatorname{sign}\left(\frac{\partial\,\phi_G^{d^{**}}}{\partial\,\hat{\phi}_G}\right) =
$$

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

$$\operatorname{sign}\left(\frac{\partial\,\sigma_G^{**}}{\partial\,\hat{\phi}_G}\right) = \tag{A.8}$$

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

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

$$\operatorname{sign}\left(\frac{\partial\,\phi_G^{d^{**}}}{\partial\,\hat{\phi}_{-G}}\right) =$$

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

$$\operatorname{sign}\left(\frac{\partial\,\sigma_G^{**}}{\partial\,\hat{\phi}_{-G}}\right) =$$

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

$$\operatorname{sign}\left(\frac{\partial\,\phi_G^{d^{**}}}{\partial\,q_G}\right) = \tag{A.9}$$

$$\operatorname{sign}\Bigg(\Bigg(\left(\left(\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}\right)^2 - \frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,{\phi^{\text{d}}_{\text{-G}}}^2}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,{\hat{\sigma}_{\text{-G}}}^2}\right)\left(\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,{\hat{\sigma}_{\text{G}}}^2}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\text{q}_{\text{G}}} - \frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat{\sigma}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\text{q}_{\text{G}}}\right) +$$

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

$$\operatorname{sign}\left(\frac{\partial\,\sigma_G^{**}}{\partial\,q_G}\right) = \tag{A.10}$$

$$\operatorname{sign}\Bigg(\left(\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,{\phi^{\text{d}}_{\text{-G}}}^2}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,{\hat{\sigma}_{\text{-G}}}^2} - \left(\frac{\partial^2\,\mathcal{L}^{**}_{\text{-G}}}{\partial\,\phi^{\text{d}}_{\text{-G}}\,\partial\,\hat{\sigma}_{\text{-G}}}\right)^2\right)\left(\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\text{q}_{\text{G}}}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\phi^{\text{d}}_{\text{G}}\,\partial\,\hat{\sigma}_{\text{G}}} - \frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,{\phi^{\text{d}}_{\text{G}}}^2}\,\frac{\partial^2\,\mathcal{L}^{**}_{\text{G}}}{\partial\,\hat{\sigma}_{\text{G}}\,\partial\,\text{q}_{\text{G}}}\right) +$$

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

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[43]. Since the sign of

[43] To see that this also holds in the last bracket–terms of equations (A.8) and (A.7), note that we obtain after 'expansion' and simplification the terms $-\frac{\partial\, s^{\phi^s_G}(\phi^{**}_{\tilde{G}},\hat{\phi}_G)}{\partial\,\phi_{\tilde{G}}}\,\frac{\partial^2\, s^{\phi^s_G}(\phi^{**}_{\tilde{G}},\hat{\phi}_G)}{\partial\,{\phi_{\tilde{G}}}^2}(1-q_G)\hat{\sigma}^{**}_G$ (the sign of which expression is negative/positive for $G = L/H$) and $\frac{\partial\, s^{\phi^s_G}(\phi^{**}_{\tilde{G}},\hat{\phi}_G)}{\partial\,\phi_{\tilde{G}}}\,\frac{\partial^2\, s^{\phi^s_G}(\phi^{**}_{\tilde{G}},\hat{\phi}_G)}{\partial\,\phi_{\tilde{G}}\,\partial\,\phi^s_{\tilde{G}}}(1-q_G)\hat{\sigma}^{**}_G$ (the sign of which expression is positive for both cultural groups). Note further that as a result of the

$\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \partial \hat{\sigma}_{\mathrm{G}}}$ and $\frac{\partial^2 \mathcal{L}^{**}_{\text{-G}}}{\partial \phi^{\mathrm{d}}_{\text{-G}} \partial \hat{\sigma}_{\text{-G}}}$ are ambiguous, this also holds for all bracket–terms above where these expressions are included. For $G = L/H$, consider the matrix

$$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 \hat{\phi}_{\mathrm{G}}}}{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \partial \hat{\phi}_{\mathrm{G}}}} & \frac{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \partial \phi^{\mathrm{s}}_{\mathrm{G}}}}{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \partial \phi^{\mathrm{s}}_{\mathrm{G}}}} & \frac{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \partial \mathrm{q}_{\mathrm{G}}}}{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \partial \mathrm{q}_{\mathrm{G}}}} & \frac{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \partial \phi^{\mathrm{d}}_{\text{-G}}}}{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \partial \phi^{\mathrm{d}}_{\text{-G}}}} \\ \frac{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \partial \phi_{\mathrm{G}}}}{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \partial \phi_{\mathrm{G}}}} & \frac{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \partial \hat{\phi}_{\mathrm{G}}}}{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \partial \hat{\phi}_{\mathrm{G}}}} & \frac{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \partial \phi^{\mathrm{s}}_{\mathrm{G}}}}{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \partial \hat{\phi}_{\mathrm{G}}}} & \frac{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \partial \mathrm{q}_{\mathrm{G}}}}{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \partial \mathrm{q}_{\mathrm{G}}}} & \frac{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \partial \phi^{\mathrm{d}}_{\text{-G}}}}{\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \partial \phi^{\mathrm{d}}_{\text{-G}}}} \end{pmatrix}'$$

and note that for cultural group L/G, all entries of this matrix are strictly negative/positive. Next, denote $b_{G1}(P) := \max/\min b_G(1)$ and $b_{G2}(P) := \max/\min b_G(2)$, where $P = \{\phi_L, \phi_H, \phi^s_L, \phi^s_H, \hat{\phi}_L, \hat{\phi}_H, q_H\}$, and let, $\forall G \in \{L, H\}$

$$\frac{\partial^2 \mathcal{L}^{**}_g}{\partial \phi^d_g \partial \hat{\sigma}_g} < / > b_{G1}(P) \frac{\partial^2 \mathcal{L}^{**}_g}{\partial {\phi^d_g}^2} \text{ and } \frac{\partial^2 \mathcal{L}^{**}_g}{\partial \phi^d_g \partial \hat{\sigma}_g} < / > b_{G2}(P) \frac{\partial^2 \mathcal{L}^{**}_g}{\partial \hat{\sigma}^2_g} \quad \text{(A.11)}$$

which basically requires that the two socialization instruments must not be too strong substitutes in the parental optimization problem (4) (or that the optimization problem must be sufficiently concave in both socialization instruments compared to the cross–concavity)[44]. Using the sign–matrices (A.4) and (A.5) it is straightforward to show that under the conditions (A.11), the comparative statics effects of Proposition 5 (a) hold.

To show part (b) of Proposition 5, let us first denote for brevity the four main bracket–terms of sign–equation (A.9) for cultural group $G \in \{L, H\}$

---

matrix–multiplications according to the Implicit Function Theorem, the fourth bracket–term of sign–equation (A.6) would actually be $\frac{\partial^2 \mathcal{L}^{**}_{\text{-G}}}{\partial \phi^{\mathrm{d}}_{\text{-G}} \partial \hat{\sigma}_{\text{-G}}} \frac{\partial^2 \mathcal{L}^{**}_{\text{-G}}}{\partial \hat{\sigma}_{\text{-G}} \partial \phi^{\mathrm{d}}_{\mathrm{G}}} - \frac{\partial^2 \mathcal{L}^{**}_{\text{-G}}}{\partial \phi^{\mathrm{d}}_{\text{-G}} \partial \phi^{\mathrm{d}}_{\mathrm{G}}} \frac{\partial^2 \mathcal{L}^{**}_{\text{-G}}}{\partial {\hat{\sigma}_{\text{-G}}}^2}$. But after 'expansion' and rearrangements of this term, we obtain the expression as stated in the equation. Finally, after 'expansion' and simplification, $\frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \partial \phi^{\mathrm{d}}_{\text{-G}}} \frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \partial \mathrm{q}_{\mathrm{G}}} - \frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \phi^{\mathrm{d}}_{\mathrm{G}} \partial \phi^{\mathrm{d}}_{\text{-G}}} \frac{\partial^2 \mathcal{L}^{**}_{\mathrm{G}}}{\partial \hat{\sigma}_{\mathrm{G}} \partial \mathrm{q}_{\mathrm{G}}}$ becomes zero. So, this term that would be present in the last comparative statics equation (A.10) drops out.

[44]The search for condition (A.11) can be simplified by analyzing the structure of the matrices b_G. First note that the fractions in each row of the matrices are inverses to each other. Second, in the 'expanded' form of these fractions, it is easy to see that there are only three different expressions in each column, namely

$$c_{G1} := \frac{\left(\frac{\partial u^{\phi_G}\left(\phi^{d^{**}}_G\right)}{\partial \phi^d_G}\right)^2 \frac{\partial^2 c\left(\hat{\sigma}^{**}_G, u^{\phi_G}\left(\phi^{d^{**}}_G\right)\right)}{\partial {\phi^d_G}^2} - \frac{\partial^2 \mathrm{u}^{\phi_G}\left(\phi^{\mathrm{d}^{**}}_{\mathrm{G}}\right)}{\partial {\phi^{\mathrm{d}}_{\mathrm{G}}}^2}\left(1 - \frac{\partial c\left(\hat{\sigma}^{**}_G, u^{\phi_G}\left(\phi^{d^{**}}_G\right)\right)}{\partial u^{\phi_G}\left(\phi^d_G\right)}\right)}{\frac{\partial u^{\phi_G}\left(\phi^{d^{**}}_G\right)}{\phi^d_G} \frac{\partial^2 c\left(\hat{\sigma}^{**}_G, u^{\phi_G}\left(\phi^{d^{**}}_G\right)\right)}{\partial \hat{\sigma}_G \partial u^{\phi_G}\left(\phi^d_G\right)}}$$

$$c_{G2} := \frac{\hat{\sigma}^{**}_G}{(1 - q_G)\left(\phi^{d^{**}}_G - \phi^{d^{**}}_{-G}\right)}$$

$$c_{G3} := \frac{\frac{\partial^2 s^{\phi^s_G}\left(\phi^{**}_{\tilde{G}}, \hat{\phi}_G\right)}{\partial {\phi_{\tilde{G}}}^2} \hat{\sigma}^{**}_G (1 - \hat{\sigma}^{**}_G)}{\frac{\partial^2 s^{\phi^s_G}\left(\phi^{**}_{\tilde{G}}, \hat{\phi}_G\right)}{\partial {\phi_{\tilde{G}}}^2}(1 - q_G)(1 - \hat{\sigma}^{**}_G)\left(\phi^{d^{**}}_G - \phi^{d^{**}}_{-G}\right) - \frac{\partial s^{\phi^s_G}\left(\phi^{**}_{\tilde{G}}, \hat{\phi}_G\right)}{\partial \phi_{\tilde{G}}}}$$

in the order of their appearance as u_{-G}, v_G, w_{-G}, x_G and the second and fourth main bracket–terms of sign–equation (A.10) as y_G, z_G (the first and third equal $-u_{-G}$ and $-w_{-G}$).

Let us consider cultural group L. Using the sign–matrices (A.4) and (A.5) as well as condition (A.11), it follows that $\text{sign}(u_H, v_L, w_H, x_L, y_L, z_L) = (-,-,-,+,-,+)$. From equation (A.9) we easily obtain $\frac{\partial \phi_L^{d^{**}}}{\partial q_L} < / = / > 0 \Leftrightarrow \frac{u_H}{w_H} < / = / > -\frac{x_L}{v_L}$; and from equation (A.10) we get $\frac{\partial \sigma_L^{**}}{\partial q_L} > / = / < 0 \Leftrightarrow \frac{u_H}{w_H} < / = / > -\frac{z_L}{y_L}$. It follows that $\left(\text{sign}\left(\frac{\partial \phi_L^{d^{**}}}{\partial q_H}\right), \text{sign}\left(\frac{\partial \sigma_L^{**}}{\partial q_H}\right)\right) \in \{(-,+),(+,-),(0,0)\}$ iff $x_L y_L - v_L z_L = 0$. After simplifications, the last equation can be shown to equal $u_L\left(\frac{\partial^2 \mathcal{L}_L^{**}}{\partial \phi_L^d \, \partial q_L}\frac{\partial^2 \mathcal{L}_L^{**}}{\partial \hat{\sigma}_L \, \partial \phi_H^d} - \frac{\partial^2 \mathcal{L}_L^{**}}{\partial \phi_L^d \, \partial \phi_H^d}\frac{\partial^2 \mathcal{L}_L^{**}}{\partial \hat{\sigma}_L \, \partial q_L}\right) = 0$. But after 'expansion' and simplifications, the last term in brackets equals zero which establishes the desired result. The proof of the analogous result for cultural group H follows the same steps using $\text{sign}(u_L, v_H, w_L, x_H, y_H, z_H) = (-,+,+,+,-,-)$. □

A.2.6 Proof of Proposition 6

1. Case $\phi_H > \phi_L$

(a) It rests to show that $\hat{\sigma}_G^* < 1$, $\forall G \in \{L, H\}$. Suppose to the contrary that any $\hat{\sigma}_G^* = 1$. Then, since in this case $\phi_{\tilde{G}}^* = \phi_G^{d^*}$, the only corresponding optimal demonstrated behavior would be $\phi_G^{d^*} = \phi_G \equiv \hat{\phi}_G$. But then, FOC (A.3) reduces to $-\frac{\partial c\left(1, u^{\phi_G}(\phi_G)\right)}{\partial \hat{\sigma}_G} - \mu_1 = 0$ which can not be fulfilled.
(b) Follows directly from Proposition 3 for the case $\hat{\phi}_G \equiv \phi_G$. □

2. Case $\phi_H = \phi_L$

Assume that any parents G choose $\hat{\sigma}_G > 0$. Then, for FOC (A.3) to hold, it is necessary that $\phi_G^d \neq \phi_{-G}^d$. Thus, at least one of the demonstrated cultures must satisfy $\phi_G^d \neq \phi_G$, $G \in \{L, H\}$. In case that for this group $\phi_G^d > (<) \phi_G$, it is needed by FOC (A.2) that $\phi_{\tilde{G}} < (>) \hat{\phi}_G = \phi_G$. For this to be possible,

---

and their inverses, respectively. Specifically, we then obtain, for $G \in \{L, H\}$

$$b_G = \begin{pmatrix} c_{G1} & c_{G2} & c_{G2} & c_{G3} & c_{G3} \\ c_{G1}^{-1} & c_{G2}^{-1} & c_{G2}^{-1} & c_{G3}^{-1} & c_{G3}^{-1} \end{pmatrix}'$$

and hence (again for $G = L/H$), $b_{G1}(P) = \max/\min\{c_{G1}, c_{G2}, c_{G3}\}$ and $b_{G2}(P) = \max/\min\{c_{G1}^{-1}, c_{G2}^{-1}, c_{G3}^{-1}\}$.

It is also interesting to observe that if only the first part of condition (A.11) would be fulfilled, then still the signs of the first row of the sign–matrix in Proposition 5 (a) would be satisfied. Only for the additional results of the second row, the full condition (A.11) is required.

it is necessary that $\phi_{-\tilde{G}} > (<)\hat{\phi}_{-G} = \phi_{-G} = \phi_G$. Thus, the two pairs of inequalities $\phi_G^d > (<)\phi_{-G}^d$ and $\phi_{\tilde{G}} < (>)\phi_{-\tilde{G}}$ must hold in the case that $\hat{\sigma}_G > 0$. But since for both cultural groups, the culture of the children is a convex combination of the representative demonstrated culture of the public social space and the parental demonstrated culture (and since at least one $\hat{\sigma}_G > 0$), this yields a contradiction (this issue has already be shown in more detail in the proof of Proposition 3, 1. (b)+(c), in Appendix A.2.3). □