

# The Dynamics of Continuous Cultural Traits in Social Networks

Berno Buechel\*      Tim Hellmann†      Michael M. Pichler‡

June 16, 2014

## Abstract

We consider an overlapping generations model where *continuous* cultural traits are transmitted from an adult generation to the children. A weighted *social network* describes how children are influenced not only by their parents but also by other role models within the society. Parents can invest into the purposeful socialization of their children by strategically displaying a cultural trait (which need not coincide with their true cultural trait). We observe a cultural substitution effect when parents choose their behavior optimally. Based on Nash equilibrium behavior, we then study the dynamics of cultural traits throughout generations. These converge if parent's influence on their children is large enough compared to the social environment's influence. Under convergent dynamics, closed subgroups fully assimilate, while heterogeneous traits prevail in the other groups. Speed of convergence is low when parents' incentives to socialize their children to the own trait are high.

Keywords: cultural transmission, continuous cultural traits, social networks, persistence of cultural traits, opinion dynamics

JEL-Classification: A14, C72, D85, D83, Z13.

## 1 Introduction

Economic behavior and outcomes are fundamentally shaped by individual value systems such as culturally transmitted preferences, attitudes, opinions, beliefs, etc. For instance, risk preferences and patience are classical determinants of economic decisions. It is even shown that inherited trust is a major factor in determining economic growth (Algan and Cahuc, 2010; Tabellini, 2010). The question of how these traits are formed and evolve is hence of central interest.

---

\*Department of Economics, University of Hamburg, Email: [berno.buechel@uni-hamburg.de](mailto:berno.buechel@uni-hamburg.de)

†Corresponding Author. Center for Mathematical Economics, Bielefeld University, P.O. Box 100131, 33501 Bielefeld, Germany. Phone: +49-521-1062573, Fax: +49-521-1062997, Email: [tim.hellmann@uni-bielefeld.de](mailto:tim.hellmann@uni-bielefeld.de).

‡Center for Mathematical Economics, Bielefeld University, Email: [mpichler@uni-bielefeld.de](mailto:mpichler@uni-bielefeld.de)

In this paper, we provide a theory of the evolution of cultural traits. In contrast to most of the literature (see Bisin and Verdier, 2010, for a comprehensive survey), we model cultural traits as a *continuous* variable rather than a discrete variable. This modeling approach better reflects the continuous nature of cultural traits such as risk preferences, patience, and trust. That these traits are rather modeled by a continuous variable becomes evident when considering e.g. Arrow-Pratt measures of risk aversion, or discount factors in  $[0, 1]$  as measures of patience.<sup>1</sup> To model the evolution of these traits across generations, we employ an overlapping generations (OLG) society. Parents care about their children’s adopted cultural trait and have a desire that this adopted trait is close to their own trait, an assumption called *imperfect empathy* in the literature.<sup>2</sup> Following empirical evidence (e.g. Dohmen et al., 2012), children are assumed to learn from observable cultural traits of their parents and of their social environment, represented by a *social network*. Surprisingly, this local aspect of trait formation has largely been ignored in the literature on formation of cultural traits even though empirical evidence suggests that the social network plays a crucial role: first, Dohmen et al. (2012) show that the degree of risk aversion and trust can not only be explained by the parent’s risk preferences and trust attitudes, but also by the level of these traits in the local social environment; second, the fact that cultural traits often differ across geographic regions, e.g. cities (Guiso et al., 2008; Voigtländer and Voth, 2012), or countries (Algan and Cahuc, 2010),<sup>3</sup> is difficult to explain without modeling a local structure. Finally, social connections also seem crucial for questions of persistence of cultural traits and of assimilation as we discuss below.

Emphasizing the social network and the socialization incentives, or in other words, the degree of imperfect empathy, we use our model to study (i) how cultural traits evolve, (ii) under which conditions heterogeneous or homogeneous societies emerge, and (iii) how long this process takes before settling down.

These questions are motivated by empirical studies that provide substantial evidence that some cultural traits are persistent throughout many generations. For the example of trust, Guiso et al. (2008) report persistence of trust levels in various Italian cities, Nunn and Wantchekon (2011) show that (mis-)trust attitudes in African families are prevailing throughout many generations resulting from slave trade history, and Voigtländer and Voth (2012) find that (mis-)trust attitudes towards the Jewish population persisted over many centuries in German cities dating back to the 14th century.<sup>4</sup> Moreover, immigrant families in the United States are shown to retain their trust attitudes throughout many generations (Algan and Cahuc, 2010) if connections to their country of origin are kept. Importantly, not only the formation of cultural traits (Dohmen et al., 2012), but also the evolution of cultural traits and the question

---

<sup>1</sup>Also trust attitudes can be measured on a continuous scale representing different intensities of trust ranging from attitudes such as “you can’t be too careful” to attitudes like “most people can be trusted”.

<sup>2</sup>Imperfect empathy means that parents care about their children, but evaluate their (the children’s) actions through the own (the parental) utility function (Bisin and Verdier, 1998). This will imply that parents want their children to become as they (the parents) are.

<sup>3</sup>Note that also within regions strong correlations of cultural traits are observed (Dohmen et al., 2012).

<sup>4</sup>Persistence is also documented for other cultural traits such as attitudes towards female labor force participation (Fernandez et al., 2004; Alesina et al., 2011), fertility (Fernandez and Fogli, 2006; Cygan-Rehm, 2012), and a preference for education (Botticini and Eckstein, 2005).

of persistence seem to be highly dependent on social connections.<sup>5</sup>

A theoretical explanation for the phenomenon of global persistence of cultural traits is provided in the seminal paper by Bisin and Verdier (2001) and the following literature (see Bisin and Verdier, 2010). They show in a population dynamics framework of dichotomous cultural traits that the imperfect empathy assumption is sufficient to generate long-run heterogeneity of cultural traits. This is due to the cultural substitution effect, i.e. that parents exert more effort to increase the probability that the child adopts the same cultural trait when the frequency of the own trait in the population is lower.

Although the model of Bisin and Verdier (2001) can be generalized to a discrete number of traits (Bisin et al., 2009), we will see that as soon as the cultural trait under consideration is modeled as a continuous variable, the observed long-term persistence of heterogeneous traits does not carry over (Remark 1). Thus, imperfect empathy alone cannot account for long-term persistence of heterogeneous cultural traits. Moreover, due to the discrete nature of traits, (partial) assimilation cannot be explained by this model. In addition, this standard model does not include local interaction.

To bridge these gaps, we complement and extend the ideas of Bisin and Verdier (2001) by (a) modeling cultural traits as a continuous variable, and (b) by assuming that children learn from parents and their friends through local interaction represented by a *social network* and not necessarily from the whole population. In our model, we adopt the assumption of *imperfect empathy* of Bisin and Verdier (2001) which leads parents to use costly controls to influence their children's traits. We consider a particular socialization instrument which is derived from social learning theory (Bandura, 1977): since children learn primarily from the observed behavior of parents (and other social contacts), we assume that the parent's primary socialization instrument is their behavior, or their socio-economic actions taken. When solving for optimal behavior of parents, we show that each adult deviates from its true trait into the opposite direction of the aggregate behavior of the environment (relative to own true trait), in order to countervail the subjectively negative influence of the environment on its child (Proposition 1). The extent of deviation is increasing with the "cultural distance" between the parent and its social environment.<sup>6</sup>

Assuming Nash equilibrium play in every generation, we then study the dynamics of cultural traits. First, we illustrate by a simple two-dynasty example (Section 3) how the cultural traits evolve. We show that cultural traits converge in the long-run to a homogeneous trait such that relative positions are persistent if children are primarily influenced by their parents. By persistence of relative positions of cultural traits, we mean that one family has the lower trait than the other at any point in time, e.g. one family is always more risk averse than the other family, although in the very long run traits become more and more homogeneous. If, however, each child is more influenced by the other family than the own, then relative positions of traits switch from one generation to the next. For strong enough incentives to socialize the children, this even

---

<sup>5</sup>For instance, Voigtländer and Voth (2012) argue that few interactions between German cities (due to lower mobility before World War I. compared to today) are one of the main factors for persistence of different attitudes towards Jews in different cities. For cities that were "well-connected" (e.g. German Hanse-cities) such persistence cannot be observed.

<sup>6</sup>Analogous effects of *cultural substitution* are crucial in providing a theoretical explanation for heterogeneous cultures in the framework of dichotomous traits (Bisin and Verdier, 2001).

leads to divergent dynamics.

When allowing for more than two dynasties (Section 4), additional steady states are possible also admitting heterogeneous cultural traits. Conditions that the dynamics reach a steady state generalize nicely from 2 to  $n$  dynasties: convergence is guaranteed if (i) the network is positive definite (Proposition 4) or if (ii) the degree of imperfect empathy is sufficiently low (Proposition 5).<sup>7</sup> Positive definiteness is implied if parents have sufficiently high influence on their children. This rather natural assumption is also empirically well supported (Dohmen et al., 2012). While these large socialization weights foster convergence to a steady state, incentives to socialize (high degree of imperfect empathy) are detrimental for convergence in two ways. First, if the socialization incentives are small enough, then convergence always obtains for any given network. Second, even if convergence is guaranteed for arbitrary socialization incentives (e.g. because of large enough socialization weights), then speed of convergence is reduced by the parents' socialization efforts (Proposition 6). This yields an explanation for the persistence of cultural traits: the higher the degree of imperfect empathy, the longer cultural traits persist (a complementary conclusion as in Bisin and Verdier, 2001).

Convergence to a steady state does not necessarily mean that we observe complete assimilation or a melting pot society. Rather, heterogeneity of cultural traits is observed even in the long-run when connections to more than one cultural group are given. Thus, persistence of immigrants' cultural traits (e.g. trust attitudes, Algan and Cahuc, 2010) may be explained by connections into both the country of origin (relatives, friends,...) and the immigrant country (as we will see in Example 2). This view is also shared by Aleksynska (2011) where a particular focus is on ties to the country of origin. In a similar way, we find a theoretical explanation for the finding that relatively disconnected cities in Germany experienced persistence of (positive or negative) attitudes towards the Jewish population, while well connected cities (e.g. Hanse-cities) rather are observed to display alternating attitudes over time (Voigtländer and Voth, 2012).

**Related Literature** Besides the well-established literature on the transmission of discrete traits, the literature on continuous traits is still small,<sup>8</sup> despite its importance for empirical applications (Bisin and Topa, 2003). Important early treatments of the topic are Cavalli-Sforza and Feldman (1981) in a theoretical, and Otto et al. (1994) in an empirical context. Cavalli-Sforza and Feldman (1981) propose a model where a child's continuous trait is formed as the linear combination of its parents' trait and the average trait in the society. Several models follow this approach and endogenize the trait formation process (for recent approaches, see Doepke and Zilibotti, 2008; Vaughan, 2010; Pichler, 2010; Panebianco, 2014).

While in Vaughan (2010) and Panebianco (2014) the parent's choice variable (socialization instrument) is the time spent with children, Pichler (2010) introduces behavioral choices of parents. The first modeling choice is fully analogous to Bisin and Verdier (2001), the latter rather reflects social learning theory (Bandura, 1977). Additionally, Vaughan (2010) models strategic interaction between children and their peers. In our paper, we instead consider strategic interaction in the adult generation and explicitly

---

<sup>7</sup>If both conditions are not satisfied, convergence is still possible, but also diverging dynamics as outlined above for two dynasties may arise.

<sup>8</sup>See Bisin and Verdier, 2010, for a comprehensive overview of both branches of literature.

model the social network. Panebianco (2014) studies cultural groups and their attitudes towards each other. Since the only variable of choice is the time spent with children, the dynamics in the model of Panebianco (2014) are given by a Markov process where convergence is easily established without restrictions on the network. Introducing behavioral choices, our model yields more complex dynamics such that the law of motion may also contain negative inter-dependencies. We also extend the approach by Pichler (2010) from global interaction to local interaction.

Doepke and Zilibotti (2008) present a model on the inter-generational transmission of continuous traits (patience and work ethos), which is used to explain the industrial revolution. In particular, they show how a generation of patient and hard working industrialists could outperform the formerly leading class of aristocrats who failed to socialize their children to these values. In contrast to other models of transmission of traits, including ours, this model includes feedback of economic conditions on the formation of traits. As a consequence it does not predict convergence of traits (to a homogeneous or heterogeneous state), but rather some kind of cycle or fluctuation since economic prosperity reduces the propensity to transmit values that are the basis of prosperity.

Another branch of literature related to our work is the literature on opinion dynamics (in social networks) introduced, among others, by DeGroot (1974) (see e.g. Jackson 2008, for a discussion). In the basic DeGroot model, individuals exchange opinions by reporting their opinions and update them according to a weighted average of other individuals' opinions. Convergence of opinions is then obtained under mild conditions on the interaction structure (strong connectedness and aperiodicity). Our work presents also a generalization of the DeGroot model such that strategic interaction in expressed opinions is introduced. In our model, strategic interaction leads to overstatement of opinions (Proposition 1), a similar, but less extreme behavior as in Kalai and Kalai (2001), where polarization of opinions is obtained. As a consequence, convergence cannot be as easily obtained as in the DeGroot model. Moreover, for the case of convergence, we show that the speed of which is reduced by introducing this kind of strategic interaction. This contrasts with a related model on opinion formation (Buechel et al., 2014) and shows that socialization investments contribute to prolong cultural heterogeneity.

## 2 Formation of Cultural Traits

### 2.1 Model

Consider an overlapping generations society which is populated by the adults of a finite set of dynasties  $N = \{1, \dots, n\}$ . At the beginning of any given period  $t \in \mathbb{N}$ , adults reproduce asexually and have exactly one offspring. This standard simplification keeps the population size constant. The traits that we consider (e.g. trust, risk preferences, and patience) are of continuous nature (e.g. different intensities of trust, Arrow-Pratt measures of risk aversion, and discount factors). We restrict the analysis to the transmission of one such continuous trait. Hence, let  $\mathcal{I} \subseteq \mathbb{R}$  be a convex compact set that contains all possible intensities/degrees of a trait. For instance, if patience is represented by a discount factor, then  $\mathcal{I} = [0, 1]$ . Each adult is characterized by a variable

$\phi_i(t) \in \mathcal{I}$ , which we call its *trait*. Taking trust as an example, low values of  $\phi_i(t)$  could then be interpreted as the  $t$ -th generation of dynasty  $i$  having low levels of trust, e.g. not trusting strangers. In a similar way, risk preferences (low values of  $\phi_i(t)$  correspond to low degrees of risk aversion), patience (low values of  $\phi_i(t)$  correspond to low discount factors), and other types of cultural traits find an interpretation in our model.

We assume that the true cultural trait is not observable. However, any adult has to make socio-economic choices. We assume that distinct cultural traits are associated with distinct socio-economic choices. Therefore, by observing the socio-economic choices of an adult, a child perceives the associated trait, which may not coincide with the adult's true trait. For instance, a parent who is very careful when interacting with strangers is displaying a low level of trust. Similarly, a parent who refuses many gambles, takes few chances and challenges in life, etc. will be perceived to be risk averse. Instead of modeling socio-economic choices explicitly, we simply assume that each adult chooses an observable cultural trait representing its socio-economic actions.<sup>9</sup> We call an adult's choice its *displayed trait* and denote it by  $\phi_i^d(t) \in \mathcal{I}$ . Let the vector  $\Phi^d(t) := (\phi_1^d(t), \dots, \phi_n^d(t))' \in \mathcal{I}^n$  collect the displayed traits of the adults. Importantly, the displayed trait  $\phi_i^d(t)$  is the choice variable of each adult and may be different from its true trait  $\phi_i(t)$ . Naturally, we will assume that deviating from the true trait is costly, see the assumption on parental utility (3).

Children are assumed to form their traits by learning from the adults' observable behavior (see also Bandura, 1977). In particular, we consider direct socialization from the parent's displayed trait  $\phi_i^d(t)$  as well as so-called oblique socialization from the child's social environment  $\phi_{N_i}^d(t)$ , i.e. representative trait of other dynasties. Let  $\sigma_{ii}$  denote the weight of the parental socialization part which determines how much the child learns from its parents (in relative terms) compared to the social environment. Factors that determine this *parental socialization weight* could include the social interaction time of the parent with its child, as well as the effort and devotion that the parent spends to socialize its child.<sup>10</sup> Then the trait formation process is given by

$$\phi_i(t+1) = \sigma_{ii}\phi_i^d(t) + (1 - \sigma_{ii})\phi_{N_i}^d(t). \quad (1)$$

Before further specifying the model, it is worth noticing the relation of Equation 1 to the trait formation processes in the literature.

**Remark 1.** *The classical model of transmission of continuous traits by Cavalli-Sforza and Feldman (1973) implies a trait formation process which can be rewritten similarly to Equation 1 (when abstracting from a noise term):  $\phi_i(t+1) = \sigma_{ii}\phi_i(t) + (1 - \sigma_{ii})\phi_N(t)$ , i.e. a child's continuous trait is formed as a combination of its parent's trait  $\phi_i(t)$  and the average trait of the whole society  $\phi_N(t)$ . For discrete traits, Bisin and Verdier (2001) assume that parents have costly controls to influence  $\sigma_{ii}$ . In a similar way, Vaughan (2010) and Panebianco (2014) endogenize the parental socialization weight  $\sigma_{ii}$  for continuous traits. However, even with the assumption of imperfect empathy (Bisin and Verdier, 1998), convergence to a homogeneous society is inevitable when traits are*

<sup>9</sup>A microeconomic foundation for this modeling assumption is given by the framework of Pichler (2010), which we employ here without explicitly modeling socio-economic choices and their relation to traits.

<sup>10</sup>See e.g. Grusec (2002) for an introductory overview of theories on determinants of parental socialization success.

continuous which is shown in Appendix A (see also Panebianco, 2014). Hence, the major result that under imperfect empathy persistence of discrete cultural traits holds does not generalize to continuous traits.

Observe that our model (cf. Equation 1) differs from the standard Cavalli-Sforza and Feldman (1973) equation above with respect to two features: First, we relax the assumption that every child is socialized by exactly the same social environment, i.e. instead of the societal average ( $\phi_N^d$ ), we consider individual averages (e.g.  $\phi_{N_i}^d$ ). Second, we distinguish between true traits  $\phi_i(t)$  and displayed traits  $\phi_i^d(t)$ . The latter is somewhat analogous to the costly controls considered in Bisin and Verdier (2001), see also Remark 2.

As mentioned in Remark 1, we do not assume that every child is socialized by exactly the same social environment, but allow them to have different socialization weights on different members of the society. Consider a  $n \times n$ -matrix  $\Sigma$ , which describes a weighted, possibly directed, social network between the dynasties. In order to account for relative influences, we assume that  $\Sigma$  is a row stochastic matrix, that is  $\sigma_{ij} \geq 0 \forall i, j \in N$ , and  $\sum_{j \in N} \sigma_{ij} = 1 \forall i \in N$ . An entry  $\sigma_{ij}$  represents the relative importance of adult  $j$  as a role model for child  $i$ . The social interaction matrix is assumed to be exogenous and its entries  $\sigma_{ij}$  can be interpreted as resulting from the relative cognitive impact (of the socialization interactions), which can be based on interaction time or on differing pre-dispositions of the children for the social learning from others. A diagonal element of  $\Sigma$  is the parent's weight  $\sigma_{ii}$  on the own child's socialization process. The elements off the diagonal  $\sigma_{ij}$  are used to average over the neighbors' displayed traits such that the representative displayed trait of child  $i$ 's social environment is given by:  $\phi_{N_i}^d(t) := \sum_{j \in N \setminus \{i\}} \frac{\sigma_{ij}}{1 - \sigma_{ii}} \phi_j^d(t)$ . These definitions allow us to restate the trait formation process (1) by  $\phi_i(t+1) = \sum_{j \in N} \sigma_{ij} \phi_j^d(t)$  for all  $i \in N$  which can be written in the following concise form for the whole society:

$$\Phi(t+1) = \Sigma \Phi^d(t). \quad (2)$$

We assume that all individuals carry over the trait that has been formed in their child period into their adult period. In the adult period then, this adopted trait  $\phi_i(t)$  guides socio-economic choices. Formally, an adult has to choose a displayed trait  $\phi_i^d(t) \in \mathcal{I}$ . We assume that this choice is evaluated with respect to two utility components: *own utility* and *inter-generational utility*, reflecting the fact that parents care about their own socio-economic choices as well as their children's adopted cultural traits. Let  $u_i : \mathcal{I} \mapsto \mathbb{R}$  represent an adult's own utility from the displayed trait  $\phi_i^d(t)$  and let  $v_i : \mathcal{I} \mapsto \mathbb{R}$  represent the utility of an adult derived from its child's adopted trait  $\phi_i(t+1)$ , i.e. the inter-generational utility component. The following specifies the assumptions on each adult's utility.

**Assumption** (Parental Utility Function). *The utility for an adult  $i \in N$ , at time  $t \in \mathbb{N}$  is given by*

$$u_i(\phi_i^d(t) | \phi_i(t)) + v_i(\phi_i(t+1) | \phi_i(t)) \quad (3)$$

with

1.  $u_i(\cdot | \phi_i(t))$  being single-peaked with peak  $\phi_i(t)$ , i.e. strictly increasing / decreasing  $\forall \phi_i^d(t) \in \mathcal{I}$  such that  $\phi_i^d(t) < / > \phi_i(t)$ ,

2.  $v_i(\cdot | \phi_i(t))$  being single-peaked with peak  $\phi_i(t)$ , i.e. strictly increasing / decreasing at all  $\phi_i(t+1) \in \mathcal{I}$  such that  $\phi_i(t+1) < / > \phi_i(t)$ ,
3.  $u_i(\cdot | \phi_i(t))$  and  $v_i(\cdot | \phi_i(t))$  being continuous, twice continuously differentiable, and strictly concave.

In part A1 we assume that own utility  $u_i$  is decreasing in the difference of the displayed cultural trait from the actual or true cultural trait. This reflects the idea that it comes with utility loss when an adult's socio-economic actions are not in line with the own (true) trait. Moreover, these dis-utilities are strictly increasing in the 'extent of the deviation'.

Part A2 postulates that inter-generational utility  $v_i$  is decreasing in the difference between the parent's trait and the trait that its child forms. There are two basic motivations to consider this case. The first one is that parents simply have an intrinsic desire that their children develop a "personality" (trait) that is as similar as possible to their own personality. The second motivation is based on a special form of parental altruism, called *imperfect empathy* (Bisin and Verdier, 1998). Parents care about the well-being of their children, but can only evaluate their child's utility under their own (not the child's) utility function – which attains its maximum at the trait of the parent.<sup>11</sup> Part A3 and the additive separability of the two utility components are technical assumptions which significantly reduce analytical complexity.

**Remark 2.** *In Bisin and Verdier (2001) parents have a general set of costly controls to influence the probability that the child learns their trait instead of learning the trait of a random individual. This "parental socialization success share" is fixed and exogenous in our baseline model and hence the socialization technology of parents is not targeted at the influence weight  $\sigma_{ii}$ .<sup>12</sup> Rather, parents can directly influence their children's trait by choice of their displayed trait (e.g. behavior or socio-economic actions) in Equation 1. While the interpretations are different, there is a natural analogy between both models: the costly choice of displayed traits in our model could well be one of the costly controls in Bisin and Verdier (2001); deviation of displayed traits in the right direction (cf. Proposition 1) influences the child's adopted continuous trait, while increasing investments into the costly controls in Bisin and Verdier (2001) influence the probabilistic analog, i.e. the probability to adopt the parents' (discrete) trait. However, in our model there is a direct externality which is not present in Bisin and Verdier (2001): by choosing a displayed trait not only the own offspring is influenced, but also other connected dynasties.*

---

<sup>11</sup>There is a form of myopia in this line of interpretation: parents do not anticipate that their children might also behaviorally deviate from their trait. Another interpretation of our model is that not only adults but also children make socio-economic choices, while parents only consider the childhood choices of their children. Young individuals (without own children) simply make socio-economic choices to maximize only the own utility part  $u$ . These choices are evaluated by their parent's inter-generational utility which is the equivalent as evaluating the associated trait. When children become adults and have their own children, the inter-generational utility component appears and to maximize their utility they will adjust their behavior since they care about their children. However, since their parents do not witness these adulthood choices anymore, this aspect does not enter their utility.

<sup>12</sup>In Section 5.1 we discuss an extension of our model, where the extend of influence  $\sigma_{ii}$  is determined as a choice variable of parents.

## 2.2 The Adults' Decisions

Summarizing the model developed in the previous section, the optimization problem of each adult  $i \in N$  becomes

$$\begin{aligned} \max_{\phi_i^d(t) \in \mathcal{I}} \quad & u_i(\phi_i^d(t) | \phi_i(t)) + v_i(\phi_i(t+1) | \phi_i(t)) \\ \text{s.t.} \quad & \phi_i(t+1) = \sigma_{ii}\phi_i^d(t) + (1 - \sigma_{ii})\phi_{N_i}^d(t), \end{aligned} \quad (4)$$

in any period  $t \in \mathbb{N}$ . The optimization problem (4) embodies the trade-off between own utility losses, resulting from choices of displayed traits that do not coincide with the true trait, and eventual improvements in the location of the child's adopted trait.

Solutions to the optimization problem are displayed traits, which are best replies to the representative environment's displayed trait  $\phi_{N_i}^d(t)$  subject to the own trait  $\phi_i(t)$ . Proposition 1 provides a characterization of the best replies, which we denote by  $\phi_i^{d*}(\phi_i(t), \phi_{N_i}^d(t))$ , and later abbreviate as  $\phi_i^{d*}(t)$ . Further let  $\partial\mathcal{I}$  be the boundary and  $\dot{\mathcal{I}}$  the interior of interval  $\mathcal{I}$ .

**Proposition 1** (Characterization of Best Replies). *For any adult  $i \in N$ , any trait  $\phi_i(t)$  and any representative trait  $\phi_{N_i}^d(t)$ , there is a unique best reply displayed trait  $\phi_i^{d*}(\phi_i(t), \phi_{N_i}^d(t))$  which satisfies the following properties:*

- (a)  $\phi_i^{d*}(\phi_i(t), \phi_{N_i}^d(t)) = \phi_i(t)$ , if one of the following conditions is satisfied:  $\sigma_{ii} \in \{0, 1\}$ ,  $\phi_{N_i}^d(t) = \phi_i(t)$ , or  $\phi_i(t) \in \partial\mathcal{I}$ .
- (b) Let  $\sigma_{ii} \in (0, 1)$ . Then  $\phi_i^{d*}(\phi_i(t), \phi_{N_i}^d(t)) < \phi_i(t)$  if and only if  $\phi_{N_i}^d(t) > \phi_i(t)$  and  $\phi_i(t) \in \dot{\mathcal{I}}$ . (Analogously for  $\phi_{N_i}^d(t) < \phi_i(t)$ ).
- (c) Let  $\sigma_{ii} \in (0, 1)$  and consider two representative traits  $\phi_{N_i}^d(t), \tilde{\phi}_{N_i}^d(t)$  such that one of the best replies  $\phi_i^{d*}(\phi_i(t), \phi_{N_i}^d(t))$  or  $\phi_i^{d*}(\phi_i(t), \tilde{\phi}_{N_i}^d(t))$  is interior. Then  $\phi_i^{d*}(\phi_i(t), \tilde{\phi}_{N_i}^d(t)) > \phi_i^{d*}(\phi_i(t), \phi_{N_i}^d(t))$  if and only if  $\tilde{\phi}_{N_i}^d(t) < \phi_{N_i}^d(t)$ .

Proposition 1 qualitatively characterizes best reply displayed traits  $\phi_i^{d*}(t)$  in terms of their deviation from the adopted cultural trait  $\phi_i(t)$ , as well as the dependence of the former on the location of the representative displayed trait in the social environment  $\phi_{N_i}^d(t)$ . Part (a) is immediate. In the polar cases where either a parent has no influence on its child, i.e.  $\sigma_{ii} = 0$ , or where the environment has no influence on the child, i.e.  $\sigma_{ii} = 1$ , a parent will not deviate from its true trait, i.e.  $\phi_i^{d*}(t) = \phi_i(t)$ . Similarly, there is no incentive to deviate from the own trait, which is the target trait for the child, when the representative social environment  $\phi_{N_i}^d(t)$  coincides with it. In these cases we thus have  $\phi_i^{d*}(t) = \phi_i(t)$ . The third case of (a) occurs when  $\phi_i(t)$  is at the boundary of the interval. Then any deviation of  $i$  would reduce both own and inter-generational utility, hence the individuals will also choose  $\phi_i^{d*}(t) = \phi_i(t)$ .

Proposition 1 part (b) shows that in the other, more generic cases, parents counteravail the respective socialization influence on their children by choosing a displayed trait that deviates from their true trait. This deviation is always into the opposite direction as the deviation of the representative displayed trait from the target trait. As an example, consider parents who attach higher importance to religion than their

local environment does. In order to avoid that their children’s value of religion is being diluted, they will behave even more religiously, e.g. go more often to church, than they would without having children. The opposite effect would obtain for secular parents in a traditional environment.

Part (c) of Proposition 1 implies that for two displayed traits in the social environment  $\tilde{\phi}_{N_i}^d(t) < \phi_{N_i}^d(t) < \phi_i(t)$ , we have  $\phi_i^{d*}(t) \left( \phi_i(t), \tilde{\phi}_{N_i}^d(t) \right) > \phi_i^{d*}(t) \left( \phi_i(t), \phi_{N_i}^d(t) \right)$ . In other words, a parent’s deviation, i.e. the distance of displayed trait from true trait, is increasing in the distance between its trait and the representative environment’s trait. To give an interpretation, we note that the distance of the chosen displayed trait from the true trait measures socialization investments since it corresponds to the own disutility a parent accepts in order to generate inter-generational utility. Thus, part (c) of Proposition 1 establishes a property of *cultural substitution* between direct and oblique socialization, showing that socialization investments are higher, the more distant (in terms of weighted average) the displayed trait in the local environment is.<sup>13</sup>

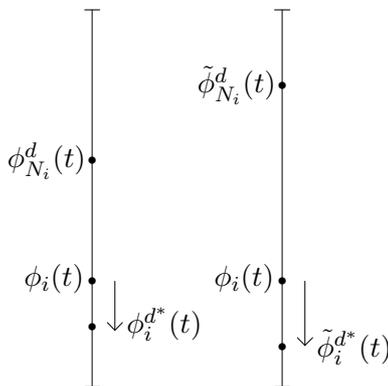


Figure 1: Characterization of Best Replies

Part (b) and part (c) of Proposition 1 are illustrated in Figure 1, where two best reply choices are depicted. Part (b) concerns the direction of the deviation of displayed trait  $\phi_i^{d*}(t)$ , respectively  $\tilde{\phi}_i^{d*}(t)$ , from true trait  $\phi_i(t)$ . Part (c) concerns the degree of deviation, which is larger in response to  $\tilde{\phi}_{N_i}^d(t)$  (on the right) than in response to  $\phi_{N_i}^d(t)$  (on the left).

We have characterized a parent’s best reply to the representative displayed trait of its social environment. To study the dynamics of cultural transmission of continuous traits, we will assume that every adult plays a best reply to the displayed trait choices of its neighbors. Hence, we assume that a Nash equilibrium is played in every period. Our previous assumptions guarantee existence of Nash equilibrium which is shown using standard techniques.

<sup>13</sup>This corresponds to the Bisin and Verdier (2001) condition of cultural substitution in a setting of dichotomous traits. That condition states that a parent’s investment into costly controls is higher the larger the share of adults in the population which have a different cultural trait. Here, we can measure the distance between own (true) trait and the weighted average displayed traits of neighbors and find that investments into behavioral deviations are higher the larger this distance.

**Proposition 2** (Nash Equilibrium Existence). *For every  $t \in \mathbb{N}$ , a Nash equilibrium in displayed trait choices exists. Denote this  $\Phi^{d^*}(t) := (\phi_1^{d^*}(t), \dots, \phi_n^{d^*}(t))'$ .*

*Proof.* In Appendix B.2.

### 3 Two Dynasties

To point out how the dynamics of cultural traits unfold and how the dynasties interact, we first discuss the case when there are only two such dynasties,  $N = \{1, 2\}$ . Consider the optimization problem given by (4) such that Assumptions A1–A3 hold. Even with this rather general functional form, the main intuition of the dynamics and the differences to the discrete traits case of Bisin and Verdier (2001) can be observed.

Let  $\phi_1(t) < \phi_2(t)$  at some point in time  $t \in \mathbb{N}$ . If the degree of risk aversion is the considered trait, then this has the interpretation that the parent in family 1 is less risk averse than the parent in family 2. By (1) the children’s adopted traits  $\phi_1(t+1)$  and  $\phi_2(t+1)$  depend on parents’ best replies of displayed traits in equilibrium such that  $\phi_1(t+1) = \sigma_{11}\phi_1^{d^*}(t) + (1 - \sigma_{11})\phi_2^{d^*}(t)$  and  $\phi_2(t+1) = (1 - \sigma_{22})\phi_1^{d^*}(t) + \sigma_{22}\phi_2^{d^*}(t)$ . From Proposition 1 we have that the displayed traits given by the best replies are more “extreme” than the true traits, i.e.  $\phi_1^{d^*}(t) \leq \phi_1(t) < \phi_2(t) \leq \phi_2^{d^*}(t)$  since the parents of generation  $t$  are investing to countervail the other family’s influence on their own child. In the context of risk, this would mean that the parent in family 1 observes that its child is also influenced by family 2 who behaves more risk averse. Hence, the parent in family 1 would take socio-economic actions which display a low degree of risk aversion (by taking gambles, etc.) since the parent wants the child to take chances in life and not to be hesitant. In similar way, the parent of family 2 would choose even less risky actions to countervail family 1’s influence.

These observations, which are based on equation (1) and Proposition 1, immediately yield the condition for child 2 to have a stronger trait intensity than child 1 such as it is the case for their parents:

$$\phi_1(t+1) \leq (<) \phi_2(t+1) \quad \Leftrightarrow \quad 1 \leq (<) \sigma_{11} + \sigma_{22}. \quad (5)$$

Thus, the relative positions of the adopted traits of the two dynasties will stay the same if and only if  $\sigma_{11} + \sigma_{22} \geq 1$ . This condition is always satisfied if children are more influenced by their own parents than by the other dynasty, i.e.  $\sigma_{ii} \geq \frac{1}{2}$  for  $i = 1, 2$ . In an extreme example, both children are only influenced by their own parents, i.e.  $\sigma_{ii} = 1$  for  $i = 1, 2$ ; then dynamics are trivial since each child becomes a copy of its parent whose displayed trait equals the true trait. Otherwise, it always holds that  $\phi_1(t+1) \in (\phi_1(t), \phi_2^{d^*}(t))$  and  $\phi_2(t+1) \in (\phi_1^{d^*}(t), \phi_2(t))$ , yielding dynamics such that the traits of the two families converge towards each other if  $\sigma_{11} + \sigma_{22} > 1$ . In that case we speak of *smooth convergence* since relative positions are maintained throughout the process. In the context of risk aversion this means that family 1 is less risk averse than family 2 at *any point in time* if children are more influenced by their own parents than by the other family. However, at each time step the families become more and more similar. Examples of smooth convergence are depicted in Figure 2a and 2b where the dynamics of true and displayed traits are shown. Hence in this two-player case

we observe assimilation and convergence to a homogeneous society although relative positions (e.g. lower degree of risk aversion of family 1) are persistent.

If instead  $\sigma_{11} + \sigma_{22} = 1$  then by (5) both families' cultural traits coincide in  $t + 1$  and hence by Proposition 1 will coincide in all following periods. Thus, after one period the dynamics of cultural traits converge, which we call *one-step convergence*. This is presented in Figure 2c. Essentially, this means that the children of both families take the same weighted average over both families' displayed trait.

Finally, if  $\sigma_{11} + \sigma_{22} < 1$  then by (5), relative positions of traits switch in each period  $t \in \mathbb{N}$ , i.e.  $\text{sign}(\phi_2(t+1) - \phi_1(t+1)) = -\text{sign}(\phi_2(t) - \phi_1(t))$  yielding alternating dynamics. These dynamics may still converge, as long as  $\sigma_{11} + \sigma_{22}$  is large enough, but they may also diverge. Figures 2d–2f display these dynamics for decreasing values of  $\sigma_{11} + \sigma_{22}$ . While the dynamics of Figures 2d and 2e converge, divergence is obtained in Figure 2f. Taking again risk aversion as an example, this means that the children (or: at least one of them) are more influenced by the other family than by the own family. Since both families exercise more extreme risk attitudes, this fact may actually lead to divergence. In our framework it may be counterintuitive to observe dynamics of cultural traits where relative positions switch each generation. However, these dynamics are induced by the hypothetical assumption that parental socialization weights are very low. In the two-dynasty case low parental socialization weights means that at least one child is more influenced by parents of the other family than by the own parents. Dohmen et al. (2012), however, present evidence that a child's cultural trait is more influenced by own parents than the social environment. Thus, the counterintuitive dynamics are induced by rather unrealistic assumptions.

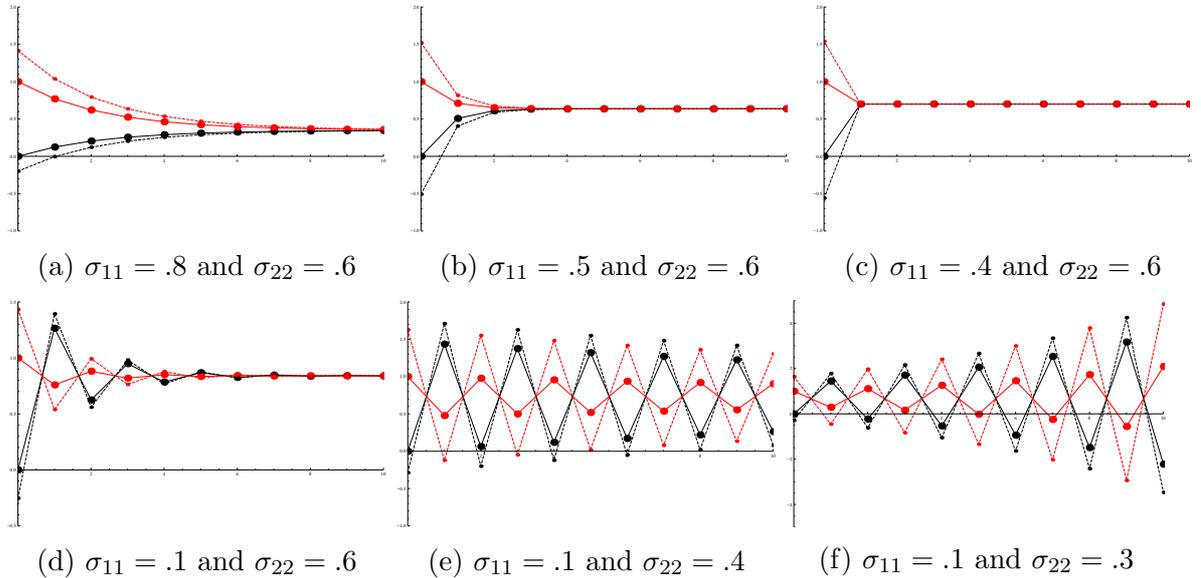


Figure 2: Dynamics of traits  $\phi_1(t)$  (solid black) and  $\phi_2(t)$  (solid red) and of displayed traits  $\phi_1^{d*}(t)$  (dashed black) and  $\phi_2^{d*}(t)$  (dashed red) for different values of parental socialization weight  $\sigma_{11}$  and  $\sigma_{22}$ .

To summarize, note that in the two-dynasty case for  $\sigma_{11} + \sigma_{22} \geq 1$ , we always obtain convergence to a melting pot society, i.e. to a homogeneous cultural trait (if  $\sigma_{11} + \sigma_{22} < 2$ ). Convergence to one cultural trait seems to be a common feature when

modeling continuous traits, see also Vaughan (2010) and Panebianco (2014). Instead, in the models of discrete traits, children adopt either one of two traits even though there may be influences from the other trait. Thus, discrete models of transmission of cultural traits are more prone to result in persistence of heterogeneous traits. Although this empirical feature is harder to obtain when modeling cultural traits by a continuous variable, this modeling approach allows us for predictions about which cultural traits will emerge; e.g. with the conditions above (large parental influence,  $\sigma_{11} + \sigma_{22} \geq 1$ ) it can be already observed that relative positions are persistent and that the society will converge to a *convex combination of initial cultural traits*. There is, however, an alternative explanation for the persistence of cultural traits: relative positions are persistent and convergence may be very slow.<sup>14</sup>

## 4 Dynamics of Cultural Traits

We now turn to the  $n$ -dynasty case, where we first characterize steady states, then derive conditions for the convergence of traits, and finally discuss persistence.

### 4.1 The Social Network and Steady States

When allowing for more than two families, it is not only the parental socialization weights that matter, but the distinct relationships  $\sigma_{ij}$  between all dynasties  $i, j \in N$  play a role.<sup>15</sup> Thus, the structure of interaction, or expressed differently, the whole network  $\Sigma$  is needed to identify convergence conditions of cultural traits and determine which groups form homogeneous traits.

It will be helpful to introduce some network specific notation. We say that there exists a connection from  $i$  to  $j$  in  $\Sigma$ , denoted by  $i \rightarrow j$ , if there exists a  $k \in \{1, \dots, n\}$  such that  $(\Sigma_{ij})^k > 0$ . Two dynasties communicate, denoted by  $i \sim j$ , if  $i \rightarrow j$  and  $j \rightarrow i$ . A dynasty  $i$  is self-communicating if  $i \rightarrow i$  which is implied by  $\sigma_{ii} > 0$ . We will assume throughout this section that every child is at least to some extent influenced by its parent, i.e.  $\sigma_{ii} > 0$  for all  $i \in N$ . Trivially,  $\sim$  defines an equivalence relation on the set  $N$  which can, hence, be partitioned into equivalence classes  $\mathcal{P}(\Sigma) = \{L_1, \dots, L_p\}$ , called self-communicating classes such that we have  $i \sim j$  if and only if there exists a  $L \in \mathcal{P}(\Sigma)$  such that  $i, j \in L$ . A communication class  $L \in \mathcal{P}(\Sigma)$  is called *essential* if for all  $i \in L$  there does not exist a  $j \notin L$  such that  $i \rightarrow j$ . A communication class is called *inessential* if it is not essential.<sup>16</sup> We will also refer to members of inessential communication classes as *rest of the world*.

Before presenting analytical results, we consider the following example to illustrate main differences to the two-dynasty case and to clarify the network specific notions introduced above.

<sup>14</sup>This explanation is also evident since there seems to be a discontinuity when  $\sigma_{11} = \sigma_{22} = 1$  where persistent heterogeneous traits emerge. To see that this is not really a discontinuity, we may decrease one  $\sigma_{ii}$  slightly by  $\epsilon$  which yields convergence to homogeneous traits, but this convergence becomes “infinitely” slow if  $\epsilon \rightarrow 0$ .

<sup>15</sup>The two-dynasty case can also be considered as a social network where all distinct relationships matter. This network, however, is fully determined by the diagonal entries of the corresponding matrix, i.e. the parents’ socialization weights, since  $\sigma_{ij} = 1 - \sigma_{ii}$ ,  $i \neq j$ .

<sup>16</sup>In network theory essential communication classes are called strongly connected and closed groups.

**Example 1.** Suppose there are seven dynasties  $N = \{1, 2, 3, 4, 5, 6, 7\}$  such that the relative influences  $\sigma_{ij}$  are given by the weighted network  $\Sigma$  defined and presented in Figure 3. In the depicted graph, a directed arc from  $i$  to  $j$  means that  $i$  is directly influenced by  $j$ , i.e.  $\sigma_{ij} > 0$ .

$$\Sigma = \begin{bmatrix} 0.8 & 0.2 & 0 & 0 & 0 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0.3 & 0 & 0 & 0.7 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0.2 & 0.4 & 0 & 0.3 \\ 0 & 0 & 0 & 0.1 & 0.1 & 0.5 & 0.3 \\ 0 & 0 & 0.3 & 0 & 0.2 & 0.1 & 0.4 \end{bmatrix}$$

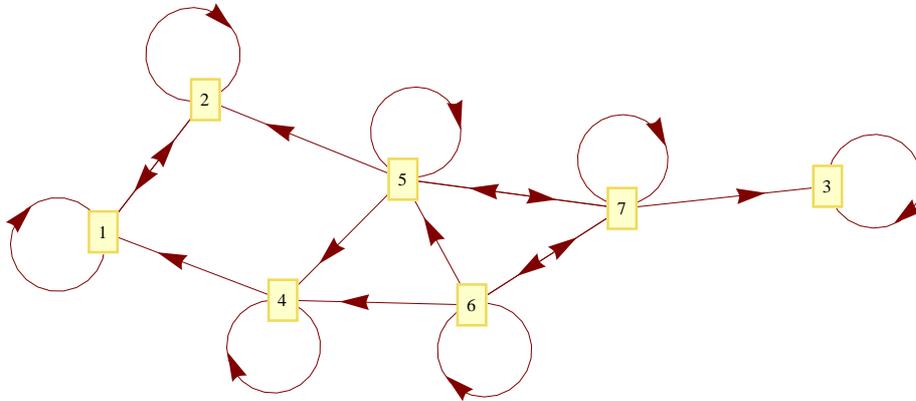


Figure 3: The network  $\Sigma$  in matrix notation and the associated graph.

The dynasties 1 and 2 are only influenced by one another, family 3 is somewhat isolated since there is no influence from other families. Family 4 is only influenced by family 1 and itself. There is a connection between all other dynasties 5, 6, and 7 and some of them are also influenced by the families 1, 2, and 3. Thus, the set of dynasties in this example can be partitioned into four communication classes  $\mathcal{P}(\Sigma) = \{L_1, L_2, L_3, L_4\}$  such that  $L_1 = \{1, 2\}$ ,  $L_2 = \{3\}$ ,  $L_3 = \{4\}$ , and  $L_4 = \{5, 6, 7\}$ .  $L_1$  and  $L_2$  are essential since they are not influenced by any other dynasty. The remaining communication classes are inessential, which we also call the rest of the world.

An example for the dynamics in this society is presented in Figure 4, where initial trait intensities are  $\Phi(0) = (110, 90, 0, 130, 20, 120, 0)'$ . The dynamics of cultural traits within essential communication classes, in this case  $L_1 = \{1, 2\}$  and  $L_2 = \{3\}$ , are independent of the dynamics of other cultural traits since those families are not influenced by any other family. In particular, smooth convergence of the cultural traits in  $L_1 = \{1, 2\}$  can be observed since  $\sigma_{11} + \sigma_{22} > 1$  (cf. Section 3). Since dynasty 3 forms a singleton essential communication class, it will keep its cultural trait  $\phi_3(t) = 0$  forever (cf. Proposition 1). Although this dynasty is isolated in some sense, it is influential for other dynasties. Within the rest of the world, i.e.  $L_3$  and  $L_4$ , convergence to heterogeneous traits can be observed. Note that in the long run, the cultural traits of dynasties in the rest of the world,  $L_3$  and  $L_4$ , are in the convex hull of the long-run cultural traits

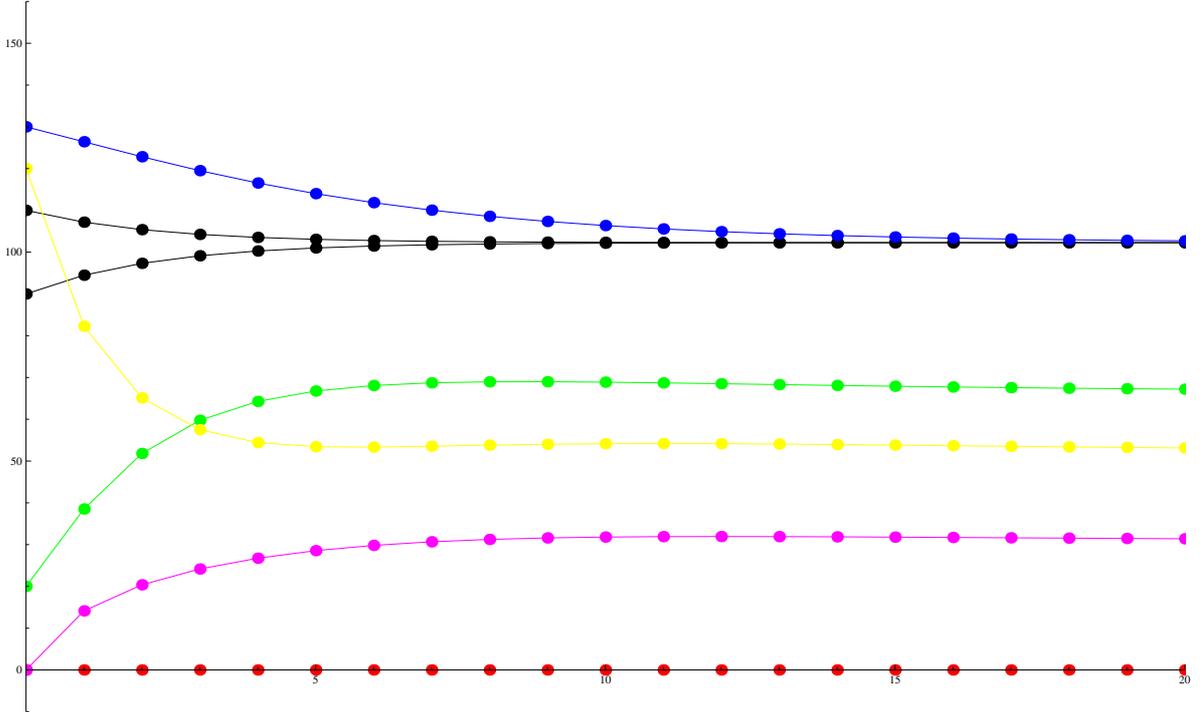


Figure 4: Dynamics of cultural traits  $\phi_1(t)$  (black),  $\phi_2(t)$  (black),  $\phi_3(t)$  (red),  $\phi_4(t)$  (blue),  $\phi_5(t)$  (green),  $\phi_6(t)$  (yellow) and  $\phi_7(t)$  (purple).

of the essential communication classes,  $L_1$  and  $L_2$ , although in period 0 this is not the case. Since dynasty 4 forms the singleton inessential communication class  $L_3$  which is only influenced by family 1, its cultural trait converges to the trait of the essential communication class  $L_1 = \{1, 2\}$ . All other families' cultural traits converge to traits which are a true and heterogeneous "mixture" of the long-run traits in the essential communication classes  $L_1$  and  $L_2$ .

Several differences to the two-dynasty case can hence be observed in Example 1: (a) the interaction structure induces a partition of the society into certain communication classes and (b) heterogeneous traits may coexist in the long run across these groups and within the rest of the world although within essential communication classes convergence to homogeneous traits is obtained.<sup>17</sup> These observations can be generalized to any steady state of the dynamics.

**Proposition 3** (Steady States). *Define a steady state as a profile of traits  $\Phi(t) \in \mathcal{I}^n$  such that  $\Phi(t+1) = \Phi(t)$ . Then the following holds in any steady state  $\Phi(t)$ :*

- (a)  $\Phi(t) = \Phi^{d^*}(t)$ , i.e. parents choose displayed traits equal to their traits.
- (b) The cultural traits of the dynasties in an essential communication class  $L \in \mathcal{P}(\Sigma)$  coincide, i.e.  $\phi_i(t) = \phi_j(t) \forall i, j \in L$ .
- (c) The cultural traits of the dynasties in an inessential communication class  $L' \in \mathcal{P}(\Sigma)$  are convex combinations of the cultural traits of the communication classes  $L \in \mathcal{P}(\Sigma)$  such that  $L' \rightarrow L$ .

<sup>17</sup>These features are well-known in the standard DeGroot model. The notable observation here is the extension of the result to our more general model and its interpretation for cultural traits.

*Proof.* In Appendix B.3.

To see that part (a) must hold, note that per definition, in any steady state, the children adopt the same cultural traits as their parents have. By (1) this implies that each parent's displayed trait must coincide with the environment's displayed trait (for  $\sigma_{ii} \in (0, 1)$ ) or that parents have either full or no influence on their child ( $\sigma_{ii} \in \{0, 1\}$ ). In each of these cases, parents behave as they are (see Proposition 1, part (a)). Hence, if two families are connected, they must share the same cultural trait in a steady state, which implies (b). Thus, (c) also holds since the inessential communication classes are influenced by the essential communication classes but not vice versa.

## 4.2 Convergence

Given this steady state description, it now remains to derive conditions under which the sequence of cultural traits actually converges to such a rest point. Even though in Example 1 the cultural traits converge, this is not always the case as shown in Section 3. To determine conditions for long-run convergence and to disentangle the effects of the social network and the effects of increasing importance of the inter-generational utility component, we will assume that both  $u_i$  and  $v_i$  are quadratic loss functions. We show in Appendix C that similar convergence conditions are required when utility is more general such that only Assumptions A1–A3 hold. Thus, let the own utility component be given by

$$u_i(\phi_i^d(t) | \phi_i(t)) := -\alpha_i (\phi_i^d(t) - \phi_i(t))^2 \quad (6)$$

and the inter-generational utility component be given by

$$v_i(\phi_i(t+1) | \phi_i(t)) = -\beta_i (\phi_i(t+1) - \phi_i(t))^2 \quad (7)$$

with  $\alpha_i, \beta_i > 0$  for all  $i \in N$ . These functional forms obviously satisfy Assumptions A1–A3. The parameter  $\alpha_i$  captures the costs of deviating from true trait. The parameter  $\beta_i$  measures the strength of *socialization incentives*, i.e. the degree of imperfect empathy (because the larger  $\beta_i$ , the more an adult  $i \in N$  wants to raise its child according to its values). These parameters are implicitly captured in the more general form of the objective function in (3). Without loss of generality, we normalize  $\alpha_i \equiv 1$  such that  $\beta_i$  represents the importance of inter-generational utility *relative to own utility*. For simplicity, we additionally assume that the space of the possible trait intensities is unbounded, i.e.  $\mathcal{I} = \mathbb{R}$ .

Then, the parents  $i \in N$  face the unrestricted optimization problems in every period  $t \in \mathbb{N}$ ,

$$\min_{\phi_i^d(t) \in \mathbb{R}} (\phi_i^d(t) - \phi_i(t))^2 + \beta_i (\phi_i(t+1) - \phi_i(t))^2 \quad (8)$$

such that  $\phi_i(t+1)$  is given by (2). With quadratic utilities, the displayed traits given by parents' best replies  $\phi_i^{d*}(t)$  are linear and can be calculated in the unique Nash equilibrium to be  $\Phi^{d*}(t) = (I + B\Sigma)^{-1} (I + B)\Phi(t)$  where  $I$  is the  $n \times n$  identity matrix and  $B$  is the diagonal matrix with entry  $\beta_i \sigma_{ii}$  in its  $i$ -th row.<sup>18</sup> Since by (2)

<sup>18</sup>Solving the first order conditions implies  $0 = \frac{\partial u_i}{\partial \phi_i^d} + \frac{\partial v_i}{\partial \phi_i^d} = (\phi_i^{d*}(t) - \phi_i(t)) - \sigma_{ii} \beta_i (\phi_i(t) - \phi_i^*(t+1))$  for all  $i \in N$ . Since in equilibrium all parents choose best replies, we have  $(I + B)\Phi(t) = B\Phi(t+1) + \Phi^{d*}(t)$ .

we have  $\Phi(t+1) = \Sigma\Phi^{d^*}(t)$ , the law of motion of the dynamics is also linear with  $\Phi(t+1) = M\Phi(t)$  where the matrix  $M$  is given by,

$$M := \Sigma(I + B\Sigma)^{-1}(I + B), \quad (9)$$

and finally

$$\Phi(t) = M^t\Phi(0).$$

Thus, in the quadratic utility case, the dynamics can be fully explained by the sequence  $M^t$ . Since conditions for convergence of such a dynamic process are well known, we can trace those back to the exogenous determinants of the matrix  $M$ . These are the social network  $\Sigma$  and the socialization incentives  $\beta_i$ . Thus, by focusing on the quadratic utility case, we are in a position to study the effects of these two parameters on the dynamics of cultural trait intensities.

In the two-dynasty case we observed that the influence weight of own parents, i.e. the diagonal of the matrix  $\Sigma$ , needs to be large enough in order to ensure convergence to a steady state. This fact generalizes to the  $n$ -dynasty case as Proposition 4 shows.

**Proposition 4** (Convergence I). *Let the parental optimization problems given by (8). Then, the following holds.*

- (a) *If  $\Sigma$  is symmetric positive definite, then for every  $\beta \in \mathbb{R}_+^n$ ,  $\lim_{t \rightarrow \infty} M^t\Phi(0)$  exists and is a steady state (for  $\Phi(0)$  arbitrary).*
- (b) *If for some eigenvalue  $\lambda$  of a non-singular  $\Sigma$  we have  $\text{Re}(\lambda) < |\lambda|^2$ ,<sup>19</sup> then there is a  $\beta \in \mathbb{R}_+^n$  such that the spectral radius of  $M$  is strictly larger than 1. Thus, for generic  $\Phi(0)$  the sequence  $\{\Phi^*(t) = M^t\Phi(0)\}_{t \rightarrow \infty}$  does not converge.*

*Proof.* In Appendix B.4.

Proposition 4 (a) shows that symmetric positive definiteness of the social network is sufficient for convergence. As the proof reveals, this condition guarantees that all eigenvalues of  $M$  are real and located in the interval  $(0, 1]$ , which implies a kind of smooth convergence (cf. Section 3). In particular, the interval formed by the convex hull of trait intensities of the current period is contained in that of the former period which drives the convergence result of Proposition 4. Therefore, all long term traits are contained in the convex hull of the initial traits  $\Phi(0)$ . For a symmetric matrix  $\Sigma$ , a sufficient condition for positive definiteness is that it is strictly diagonally dominant, i.e.  $\forall i \in N : \sigma_{ii} > \frac{1}{2}$ . Hence, we get a direct generalization of the two-dynasty result: if parents' influence on their children is larger than the environment's (together with symmetric interaction weights), then the dynamics converges to a steady state. In other words, the dynamics of cultural traits converge as long as the parental socialization weights  $\sigma_{ii}$  are large enough. Instances of non-convergence, as seen in Section 3, always

---

That  $I + B\Sigma$  is invertible can be ensured by assuming e.g. that  $\Sigma$  is symmetric positive definite. To see this note that  $I + B\Sigma$  and  $I + B^{\frac{1}{2}}\Sigma B^{\frac{1}{2}}$  have the same eigenvalues. Now if  $\Sigma$  is symmetric,  $B^{\frac{1}{2}}\Sigma B^{\frac{1}{2}}$  is also symmetric and if  $\Sigma$  is positive semidefinite and  $B \geq 0$ ,  $B^{\frac{1}{2}}\Sigma B^{\frac{1}{2}}$  is positive semidefinite. Thus,  $B\Sigma$  has non-negative and real eigenvalues which implies that all eigenvalues of  $I + B\Sigma$  are non-zero, thus  $I + B\Sigma$  is invertible. Substituting  $\Phi(t+1) = \Sigma\Phi^{d^*}(t)$  implies the result.

<sup>19</sup> $\text{Re}(\lambda)$  means the real part of eigenvalue  $\lambda$ .

go along with alternating dynamics, a fact that requires (when socialization weights are symmetric) low influence of own parents, an unrealistic assumption.

The idea that high parental influence and symmetric interaction weights ensure convergence can also be generalized to more generic functional forms of utility such that our Assumptions A1–A3 hold. However, one additional condition on the social network is required: no individual dynasty should have excessive influence on others (see Appendix C, Definition 1(iii)). This robustness result is presented in Appendix C.

Part (b) of Proposition 4 addresses matrices that are not symmetric positive definite and states the following necessary condition for convergence (subject to any  $\beta$ ):  $Re(\lambda) \geq |\lambda|^2$  for any eigenvalue  $\lambda$  of the matrix  $\Sigma$ , i.e. the real part of each eigenvalue is larger than the squared absolute value of this eigenvalue. To see how both conditions (necessary and sufficient) relate, let  $\Sigma$  be symmetric. Then, the property  $Re(\lambda) \geq |\lambda|^2$  simplifies to  $\lambda \geq \lambda^2$  and thus to  $\lambda \in [0, 1]$  since symmetric matrices only have real eigenvalues. Thus, for symmetric and non-singular social networks  $\Sigma$  both conditions coincide. Therefore, we get the following corollary for symmetric and non-singular interaction structures  $\Sigma$ : *Cultural traits converge if and only if  $\Sigma$  is positive definite.*

In the proof of the necessity part of Proposition 4 it is used that large degrees of imperfect empathy  $\beta_i$  will lead to divergence if the eigenvalue condition is not satisfied. Hence, if the inter-generational utility part receives high weight  $\beta_i$  in the utility function, then this is detrimental for convergence. As has been mentioned above, the present special case of quadratic utilities is basically a transformation of the DeGroot model. Given that convergence is satisfied in the latter for all aperiodic  $\Sigma$ , it is intuitive that we also obtain convergence if the transformation to  $M$  (as induced by the parental socialization incentives, which are embodied in the  $\beta_i$ 's) is small enough. This is confirmed as follows.

**Proposition 5** (Convergence II). *Let the parental optimization problems be as in (8). Then, for every irreducible  $\Sigma$  with strictly positive diagonal, there exists a nonempty neighborhood  $N(\mathbf{0}|\Sigma) \subset \mathbb{R}_+^n$ ,<sup>20</sup> such that  $\forall \beta \in N(\mathbf{0}|\Sigma) \cup \mathbf{0}$ , cultural traits in the society  $\Phi(t)$  converge (for  $\Phi(0)$  arbitrary).*

*Proof.* In Appendix B.5.

In the proof of this Proposition, we show first that if  $\Sigma$  has a strictly positive diagonal, then it has a simple Perron–Frobenius eigenvalue of 1 where the absolute value of all other eigenvalues is located in the interval  $(0, 1)$ . Now, the eigenvalues are continuous in the underlying matrices. Thus, it must be possible to at least slightly perturb  $\Sigma$  such that the resulting matrix  $M$  also has a unique eigenvalue 1 with the absolute value of all other eigenvalues in the interval  $(0, 1)$ . Hence,  $M^t$  converges. Notably, this holds even though  $M$  might have negative entries.

To summarize, convergence of cultural traits is implied either by small socialization incentives  $\beta_i$  or by special structures of the matrix  $\Sigma$ , which means here that socialization weights are symmetric  $\sigma_{ij} = \sigma_{ji}$  and parents' socialization weights  $\sigma_{ii}$  are large enough. Conversely, it takes both a different network structure (implied by unrealistic assumptions) and strong enough socialization incentives to obtain non-convergence.

---

<sup>20</sup> $N(\mathbf{0}|\Sigma)$  means that the size of the neighborhood around  $\beta = \mathbf{0}$  depends on  $\Sigma$ .

### 4.3 Persistence of Cultural Traits

In our model, convergence to a steady state implies that dynasties in essential communication classes converge to the same trait. Hence, the empirically observed phenomenon of persistence of cultural traits can be represented in our model in two ways.

The first has been indicated in Example 1. Suppose that most families belong to the rest of the world. Then as shown in Example 1 those families end up with heterogeneous traits in the long-run even if they belong to one communication class, which is inessential. The following example illustrates how this argument applies to the case of immigration.

**Example 2** (Immigration). *Consider a set of immigrants who have moved from some country  $B$  to another country  $A$ . Within both countries there is interaction between residents. Moreover, immigrants keep some ties to their origin country. As an example, suppose the network structure  $\Sigma$  is such that  $\mathcal{P}(\Sigma) = \{L_1, L_2, L_3, L_4\}$  where  $L_1$  and  $L_2$  are essential communication classes from country  $A$  and  $B$ , respectively. We interpret these as the leading culture of each country. For instance, these communication classes may consist of a group of dynasties who are only influenced by the cultural trait of their respective home country.  $L_3$  and  $L_4$  are inessential communication classes of residents in Country  $A$  such that the immigrants are represented by  $I \subset L_4$ . In period  $t = 0$  the immigrants arrive in Country  $A$  with a cultural trait close to their home country's leading cultural trait and find some interaction with natives from Country  $A$  in  $L_4$  and other immigrants  $I \subset L_4$ .*

*One instance of such a network is given in Example 1. In this example, let there be one immigrant represented by dynasty 7. The immigrant's home country  $B$ , which is given by a single representative dynasty 3 in Example 1, has a leading cultural trait intensity of 0. In Country  $A$ , the leading culture is represented by the two dynasties in the essential communication class  $L_1$  with a cultural trait intensity close to 100. From Example 1 it is straightforward to see that the immigrating dynasty 7 assimilates over time, but differences in traits are persistent, even in the long run (cf. Figure 4). This is due to the fact that in this example, influence by the immigrant's origin, country  $B$ , is still present. Aleksynska (2011) reports a similar phenomenon when immigrants maintain contacts to the origin country. Moreover, as it can also be seen from Figure 4, this fact also leads several natives to adopt a different cultural trait such that multiple heterogeneous traits persist in the long run.*

Thus, the first explanation for persistence of cultural traits in our model is based on heterogeneous network connections. The second explanation for persistence of cultural traits was indicated in the discussion of the two-dynasty case. Even if the society converges to a homogeneous trait intensity, relative positions of cultural trait intensities are prevailing, although the difference vanishes over time. This process might be sufficiently slow to conclude empirically that heterogeneous traits persist.<sup>21</sup> Note that in Bisin and Verdier (2001), the assumption of imperfect empathy ensures persistence of heterogeneous cultural traits in the long run. In our model of continuous traits, the assumption of imperfect empathy may have an effect for the speed of convergence.

---

<sup>21</sup>Such an interpretation of persistence is recently also adopted in empirical literature (see e.g. Giavazzi et al., 2014). In fact, Giavazzi et al. (2014) present evidence that assimilation of immigrant traits is observed, but speed of convergence depends on the trait under consideration.

As a benchmark, consider the case where parents do not care about the cultural trait of their children, i.e. where  $\beta_i = 0$  for every parent  $i \in N$ . Mathematically, the dynamics in the benchmark case are governed by the power series of  $\Sigma^t$ , while the dynamics of our model with quadratic utilities are governed by  $M^t$ , since  $\Phi(t+1) = M^t\Phi(t)$  and  $M = \Sigma$  if  $\beta_i = 0$  for any  $i \in N$ . Thus, the spectral properties of both matrices not only determine convergence conditions but also determine the speed of convergence. Since both matrices  $\Sigma$  and  $M$  have 1 as the largest eigenvalue (see Proposition 4), convergence speed is governed by the second largest eigenvalue. Let the eigenvalues of  $\Sigma$  and  $M$  be ordered according to size, i.e.  $|\lambda_1(\Sigma)| > |\lambda_2(\Sigma)| \geq \dots \geq |\lambda_K(\Sigma)|$  and  $|\lambda_1(M)| > |\lambda_2(M)| \geq \dots \geq |\lambda_K(M)|$ , such that multiple eigenvalues may occur.<sup>22</sup> Then convergence of  $M^t$  is slower than convergence of  $\Sigma^t$  if  $|\lambda_2(M)| > |\lambda_2(\Sigma)|$ , which indeed holds, as is established by the following proposition.

**Proposition 6** (Speed of Convergence). *Let the parental optimization problems be as in (8). If  $\Sigma$  is symmetric positive definite and  $\beta_i > 0$  for all  $i \in N$ , then the eigenvalues of  $M$  (which are real and positive) satisfy:  $\lambda_k(M) > \lambda_k(\Sigma)$  for all  $2 \leq k \leq K$ . Thus, socialization incentives  $\beta$  reduce the speed at which traits  $\Phi(t)$  converge for  $t \rightarrow \infty$ . Moreover,  $\lambda_k(M) \rightarrow 1$  for all  $k = 1, \dots, K$  if  $\beta_i \rightarrow \infty$  for all  $i \in N$ . That is, for large socialization incentives  $\beta$  convergence of traits becomes arbitrarily slow.*

*Proof.* In Appendix B.6.

To interpret Proposition 6, note that zero socialization incentives, i.e.  $\beta_i = 0$  for all  $i \in N$ , imply that displayed traits always coincide with true traits. Thus, Proposition 6 first shows that parents' behavioral deviations from true trait slow down convergence. Second, if the socialization incentives  $\beta_i$  grow for all  $i \in N$ , then the eigenvalues of  $M$  approach 1. This means that all families invest more and more to keep their cultural trait and the change of cultural traits from one generation to the next becomes arbitrarily small. Thus, an arbitrarily low speed of convergence obtains.

To summarize, we find that high parents' influence on own children  $\sigma_{ii}$  facilitates convergence since alternating dynamics are avoided, while higher values of socialization incentives  $\beta_i$  lead to higher behavioral overshooting and hence to slower convergence. In case of convergence, families within essential communication classes reach the same cultural trait in the long run, while heterogeneity is obtained across essential communication classes and within the rest of the world. In any case of convergence, the long-run traits are contained in the interval of the initial cultural traits.

## 5 Discussion

### 5.1 Possible Extensions

Throughout the analysis we assumed that the social network does not change over time. This assumption, which is also used in most models on opinion dynamics, is made for analytical tractability and yields a benchmark case. While, trivially, our convergence

---

<sup>22</sup>Note that for irreducible matrices  $\Sigma$  and  $M$  there is always a unique largest eigenvalue. With our assumptions in Proposition 6 we even show that all eigenvalues of both  $\Sigma$  and  $M$  are real and positive, and thus all inequalities are strict and the eigenvalues equal their absolute values.

results are robust to small vanishing perturbations on the interaction structure, it would be interesting to study the dynamics of traits when the network structure itself is endogenous. In this section, we want to discuss briefly how our model may be extended in this direction.

First, suppose that parents do not only choose their displayed trait, but are also able to control how much they influence their own children. Higher influence can be achieved by e.g. spending more time with their children, thereby limiting interactions with other members of the society (extreme cases could include home-schooling of children). Increasing the influence on own children requires higher efforts and is, hence, more costly. These ideas are in line with the literature on cultural traits (cf. e.g. Bisin and Verdier, 2001) where the socialization efforts of parents directly translate into a higher probability of learning the own trait. These socialization efforts do not imply a direct externality on others as in the case of choosing behavior.

If the time devoted to children is the only variable of choice, then this leads to trivial dynamics as noted in Remark 1 (cf. also Panebianco, 2014). Hence, as an extension of our model we briefly outline a model where parents choose their behavior *and* their time spent with children as socialization instruments while keeping the relative interaction with others unchanged. In other words, when being influenced more by their parents than before, the influence of other families in the society decreases proportionally.

Thus, as in the baseline model, let the network be exogenously given by  $\Sigma$ . In each period  $t \in \mathbb{N}$  parents  $i \in N$  may choose efforts  $x_i(t) \geq 0$  to increase the influence on its child from the given socialization endowments  $\sigma_{ii}$  to  $\frac{\sigma_{ii} + x_i(t)}{1 + x_i(t)}$ .<sup>23</sup> This modeling choice of change in influence is concave in efforts  $x_i(t)$  and approaches 1 for arbitrarily large efforts. The influences of other families  $j \neq i$  decrease proportionally and are hence given by  $\frac{\sigma_{ij}}{1 + x_i(t)}$ . Then, analogously to our baseline model (see Equation 2), children would learn from the displayed traits of others according to the adjusted learning weights such that the adopted trait will result from:

$$\phi_i(t+1) = \frac{\sigma_{ii} + x_i(t)}{1 + x_i(t)} \phi_i^d(t) + \sum_{j \neq i} \frac{\sigma_{ij}}{1 + x_i(t)} \phi_j^d(t). \quad (10)$$

Considering, as before, quadratic utility, and, moreover, quadratic cost of effort, this yields the following optimization problem of parents:

$$\min_{(\phi_i^d(t), x_i(t)) \in \mathbb{R} \times \mathbb{R}_+} (\phi_i^d(t) - \phi_i(t))^2 + \beta_i (\phi_i(t+1) - \phi_i(t))^2 + \gamma_i (x_i(t))^2 \quad (11)$$

such that  $\phi_i(t+1)$  is given by (10). As before,  $\beta_i$  denotes the degree of imperfect empathy, while  $\gamma_i$  denotes the cost of effort (both relative to the cost of behavioral deviation which is normalized to 1).

Analogously to the analysis presented in this paper, the properties of Proposition 1 still hold in equilibrium, denoted by  $(\Phi^{d^*}(t), x^*(t))$ . Additionally, we always have  $x_i^*(t) > 0$  if  $\phi_i(t) \neq \phi_{N_i}^{d^*}$  and  $\sigma_{ii} < 1$  (similar to the reasoning in Proposition 1). Note here that with the adjusted weights, the representative displayed trait  $\phi_{N_i}^{d^*} = \sum_{j \neq i} \frac{\sigma_{ij}}{(1+x_i^*) \left(1 - \frac{\sigma_{ii} + x_i^*}{1+x_i^*}\right)} \phi_j^{d^*} = \sum_{j \neq i} \frac{\sigma_{ij}}{1-\sigma_{ii}} \phi_j^{d^*}$  is still defined as before.

<sup>23</sup>As in the literature (see also Appendix A), it is also possible to assume that spending no effort implies no influence of children, i.e. socialization endowments are given by  $\sigma_{ii} = 0$ . Here we present a slightly more general approach.

Now, let  $\phi_i(t) \neq \phi_{N_i}^{d^*}$ . The FOCs imply

$$\frac{(\phi_i^{d^*}(t) - \phi_i(t))}{(\phi_i(t) - \phi_{N_i}^{d^*}(t))} = \frac{\beta_i(\sigma_{ii} + x_i^*(t))(1 - \sigma_{ii})}{(1 + x_i^*(t))^2 + \beta_i(\sigma_{ii} + x_i^*(t))^2}. \quad (12)$$

The right-hand side of (12) is decreasing in  $x_i^*(t)$  for all  $x_i^*(t) \in \mathbb{R}_+$  and all  $1/2 \leq \sigma_{ii}$ . This implies that, faced with the same  $\phi_{N_i}^{d^*}$  as in the baseline model, each family will choose a less extreme behavioral deviation than before.

To sketch the resulting dynamics briefly, let us consider the case of two dynasties with sufficiently large parental socialization endowments,  $\sigma_{ii} \geq 1/2$ . To compare these dynamics to our baseline model, we denote the original choices and outcomes by a hat, i.e.  $\hat{\phi}_i^{d^*}(t)$  represents the best reply displayed trait of player  $i$ , when this is the only variable of choice, i.e.  $x_i(t) = 0$ . As in the two-dynasty case in Section 3, let  $\phi_1(t) < \phi_2(t)$ . By the reasoning above, we have again that the best replies in displayed traits are such that  $\phi_1^{d^*}(t) < \phi_1(t) < \phi_2(t) < \phi_2^{d^*}(t)$ . From (12) it is then immediate that for  $\sigma_{11}, \sigma_{22} \geq 1/2$ , we get  $\hat{\phi}_1^{d^*}(t) < \phi_1^{d^*}(t)$  and  $\phi_2^{d^*}(t) < \hat{\phi}_2^{d^*}(t)$  which implies  $\phi_1(t) < \phi_1(t+1) < \phi_2(t+1) < \phi_2(t)$ . Thus relative positions are preserved, implying smooth convergence. Further, we also have  $|\phi_i(t+1) - \phi_i(t)| < |\hat{\phi}_i(t+1) - \phi_i(t)|$ , i.e. parents will achieve that their children's traits are closer to their own than in the original model. This holds since each dynasty is faced with a less extreme behavior of the other dynasty than in the original model, coupled with the fact that dynasties have a richer choice set.

To sum up, by introducing direct socialization efforts into our model, we still get convergence when parents are the primary socialization source, but convergence obtains at a lower speed. We conjecture that these observations also extend to the  $n$ -dynasty case for the following reasons: first, convergence is reinforced since behavior is less extreme and the diagonal is larger (recall that for all positive definite interaction structures we established convergence). Second, we expect convergence to be slower for the exact same reason, i.e. parents invest more into socialization since they have an additional socialization instrument available. Thus, persistence of cultural traits in relative positions and in terms of slow convergence is strengthened when parents can control their influence on their children. Moreover, being the primary socialization source, may be achieved not only by the requirement that the socialization endowment  $\sigma_{ii}$  is large enough, but also by low enough effort costs. Both these conditions ensure the parental influence to be large enough *in equilibrium* which may result in less demanding convergence conditions on the interaction structure. Technically, however, the dynamics result in a time-inhomogeneous product of matrices which may contain also negative entries. In Appendix C, we show how to handle such time-inhomogeneous law of motions. The assumptions on the interaction structure (requiring a symmetric ultrametric structure) are, however, more demanding than necessary when applied to the setting of endogenous efforts. Hence, we leave the details of this analysis to future research.

Another very interesting approach of introducing a time-varying interaction structure is to allow parents to control the whole network of their children. This could be achieved by moving to a certain neighborhood, or choosing a particular school for their children. Compared to determining the own influence, these network choices, however, have also implications for the following generations' networks, since it seems reasonable to assume that the network is passed over to the next generation. Deviating from the

inherited network should be costly depending e.g. on some distance measure. Under these assumptions a co-evolution of network and traits can be studied which may as well yield separation and persistence of heterogeneous traits in the long-run. While analytically challenging, such a type of model deserves extensive investigation in future research.

Moreover, we assumed throughout the present paper that the adopted trait has no effect on an individual's wealth. If such effects are considered, then additional incentives arise which feed back to the trait formation process (see, e.g., Doepke and Zilibotti, 2008). Hence, it would be ultimately desirable to understand the co-evolution of cultural traits, individual wealth, and the social network.

## 5.2 Conclusion

In this paper, we introduce a model of cultural transmission of continuous traits within a finite population. Interaction ties are captured by a social network structure. In the related literature on cultural transmission of traits, usually a continuous player set is assumed and interaction itself is global (see for a survey, Bisin and Verdier, 2010). However, empirical evidence strongly suggests that the transmission of cultural traits is local (e.g. Dohmen et al., 2012; Voigtländer and Voth, 2012). We show in this paper that not only the *socialization incentives*, but also the *interaction structure* matters for the question of whether a homogeneous society is observable in the long run.

In case of symmetric social networks, positive definiteness of the network is necessary and sufficient for convergence of continuous cultural traits (Proposition 4). This condition has a quite intuitive interpretation: if children are more influenced by their parents than by the social environment, then the dynamics converge. We have shown this when both utility components are quadratic, but similar conditions are required for general utility (see Proposition C1 in Appendix C.1).

While a deterministic model on continuous cultural traits is more likely to result in convergence to a melting pot society (cf. also Vaughan, 2010 and Panebianco, 2014) than a probabilistic model on discrete traits (e.g. Bisin and Verdier, 2001), our exercise yields two possible answers to the puzzle of the long-term persistence of heterogeneous cultural traits. First, convergence to a steady state in our model does not imply homogeneity of traits in the whole society but only within closed subgroups of it (Proposition 3). In particular, across these groups, and more interestingly, outside these groups (i.e. in the rest of the world), heterogeneous cultural traits generically coexist in the long-run. As an example, we briefly outlined how such a model can be used to explain persistence of cultural traits after immigration (Example 2). Second, speed of convergence depends on the interaction structure and, most importantly, it is reduced by the parents' socialization efforts (Proposition 6). Thus, for high socialization incentives convergence may be very slow, thereby matching empirical results of persistence of cultural traits. Indeed, Giavazzi et al. (2014) interpret persistence of cultural traits as slow convergence and present empirical evidence that the speed of evolution depends on the trait in question. In the light of our model, it is the degree of the socialization incentives for the considered trait that drives the observed speed of convergence.

Interestingly, our model is also very close to that of opinion formation dynamics of DeGroot (1974) and the succeeding literature (e.g. DeMarzo et al., 2003; Golub and Jackson, 2010). The opinion dynamics have been studied so far only with respect to

truth telling, omitting the possibility of exaggerating as strategic choice in discussion. Interpreting our model in this way, we show that the introduction of strategic interaction leads to cases of non-convergence, while the opinion dynamics in DeGroot (1974) converge whenever the matrix that represents the network has a strictly positive diagonal. Hence, the conditions for convergence that we identify require more structure on the underlying network.

## Acknowledgments

*We thank seminar participants at meetings of the Coalition Theory Network, the European Economic Association, the North American Econometric Society, the Royal Economic Society, the Society for Advancements in Economic Theory, and at the Workshop on Trust and Cultural Evolution in Valencia, the European Workshop on General Equilibrium Theory, the Workshop on Networks and Institutions in Ankara, and the World Congress of the Game Theory Society for helpful comments and discussions. We also owe gratitude to Herbert Dawid, Gerd Muehlheusser, Stefan Klößner, Christoph Kuzmics, Alexander Perli, Guido Tabellini, Jonas Tölle, and Fernando Vega-Redondo for valuable comments, and acknowledge financial support by the German Research Foundation (DFG). The paper also benefited greatly from comments by an editor and two anonymous referees.*

## Appendix

### A A more Standard Model

Here, we briefly study an alternative model, which is closer to the models in the literature (Bisin and Verdier, 2001; Vaughan, 2010; Panebianco, 2014), in particular, with respect to the socialization instrument. For this purpose, we reconsider Equation (1) and make two simplifying assumptions. First, suppose that for each parent displayed trait equals true trait, i.e.  $\phi_i^d(t) = \phi_i(t) \forall i, t$ . Second, let the relevant social environment of every child be the unweighted average of the society  $\phi_N(t)$ . Then Equation (1) reads as follows:

$$\phi_i(t+1) = \sigma_{ii}\phi_i(t) + (1 - \sigma_{ii})\phi_N(t). \quad (\text{A.1})$$

We follow the literature by assuming that parents can invest into their socialization weight  $\sigma_{ii}$ , while there are socialization costs  $C(\sigma_{ii})$ . As in Bisin and Verdier (2001), Vaughan (2010), and Panebianco (2014) we consider the case of imperfect empathy and quadratic disutilities. Then the parental optimization problem is:  $\min_{\sigma_{ii} \in [0,1]} [\phi_i(t+1) - \phi_i(t)]^2 + c\sigma_{ii}^2$ , where  $c > 0$  is a cost parameter and  $\phi_i(t+1)$  is determined by Equation (A.1).

The first order condition yields the following optimal socialization weight:

$$\sigma_{ii}^* = \frac{(\phi_i(t) - \phi_N(t))^2}{(\phi_i(t) - \phi_N(t))^2 + c} = 1 - \frac{c}{c + (\phi_i(t) - \phi_N(t))^2}. \quad (\text{A.2})$$

Observe that optimal parental weight is decreasing in the cost parameter  $c$  and increasing in the difference between a parent's trait and the societal average trait. Thus,

the more distant a parent's trait from the average trait, the higher its socialization investment – an effect which is referred to as *cultural substitution*. However, we always obtain  $\sigma_{ii}^* \in (0, 1)$ , which implies that all dynasties converge to one single trait since traits move towards the average trait. Thus, cultural substitution is not sufficient to avoid homogeneity of traits in the long run. The model we study in this paper (Equation 1) yields different dynamics: It does not guarantee convergence to a steady state (due to a different socialization instrument); and, under convergence, multiple cultural traits can emerge (due to local interaction structures).

## B Proofs

### B.1 Proof of Proposition 1

Given  $\phi_i(t)$ ,  $\sigma_{ii}$  and  $\phi_{N_i}^d(t)$ , let us define  $f : \mathbb{R} \mapsto \mathbb{R}$  s.t.  $f(x) = v_i(\sigma_{ii}x + (1 - \sigma_{ii})\phi_{N_i}^d(t))$ , i.e.  $f$  evaluates the inter-generational utility at each displayed trait of a parent. Then, a parent's maximization problem (4) can be written as follows

$$\max_{x \in \mathcal{I}} u_i(x) + f(x). \quad (\text{B.1})$$

Note  $f''(x) = v''(\sigma_{ii}x + (1 - \sigma_{ii})\phi_{N_i}^d(t))\sigma_{ii}^2$ . By the assumptions that  $u_i$  and  $v_i$  are strictly concave, we have  $u_i''(x) + f''(x) < 0$  for all  $x \in \mathcal{I}$ . Since we maximize a continuous and strictly concave function over a compact set, there is a unique solution, which we denote by  $x^*$ .

#### Proof of part (a)

1. If  $\sigma_{ii} = 0$ , then  $f(x)$  is constant for all  $x$ . Hence,  $x^* = \arg \max u_i(x) = \phi_i(t)$ .
2. If  $\sigma_{ii} = 1$ , then  $f(x) = v(x)$  such that  $x^* = \phi_i(t)$  maximizes both  $u_i(x)$  and  $f(x)$ .
3. If  $\phi_{N_i}^d(t) = \phi_i(t)$ , then again  $x^* = \phi_i(t)$  maximizes both  $u_i(x)$  and  $f(x)$  (since it implies that  $\phi_i(t+1) = \phi_i(t)$ ).
4. Let  $\phi_i(t) = \min \mathcal{I}$ . Then  $\phi_i(t+1) \geq \phi_i(t)$  for any  $x$ . Thus,  $f(x)$  is non-increasing in  $x$ , while  $u_i(x)$  is decreasing in  $x$  attaining its maximum at  $x^* = \phi_i(t)$ . And analogously, for  $\phi_i(t) = \max \mathcal{I}$ .

**Proof of part (b)** We first assume that  $\phi_{N_i}^d(t) > \phi_i(t)$  and  $\phi_i(t) \in \overset{\circ}{\mathcal{I}}$  and show that it implies  $\phi_i^{d*}(\phi_i(t), \phi_{N_i}^d(t)) < \phi_i(t)$ . Using the definition of  $f$  as above, we note that for  $\sigma_{ii} \in (0, 1)$ ,  $f(x)$  is strictly decreasing for all  $x \geq \phi_i(t)$ .  $u_i(x)$  is non-increasing such that  $x > \phi_i(t)$  cannot be optimal. At  $x = \phi_i(t)$  we have  $u_i'(\phi_i(t)) = 0$  and  $f'(\phi_i(t)) < 0$  such that  $x = \phi_i(t) - \epsilon$  is an improvement for small enough  $\epsilon > 0$ . Moreover, for small enough  $\epsilon > 0$ , this choice is interior by the assumption that  $\phi_i(t)$  is interior. And analogously for  $\phi_{N_i}^d(t) < \phi_i(t)$ .

For the opposite direction, let us assume that  $x^* < \phi_i(t)$  and suppose that it does not imply that  $\phi_i(t)$  is interior and that  $\phi_{N_i}^d(t) > \phi_i(t)$ . If  $\phi_i(t)$  is at the boundary, part (a) of Proposition 1 implies that  $x^* = \phi_i(t)$ , which contradicts the assumption. If  $\phi_{N_i}^d(t) \leq \phi_i(t)$ , then for  $x < \phi_i(t)$  both  $f(x)$  and  $u_i(x)$  are increasing such that  $x < \phi_i(t)$  cannot be optimal. A contradiction.

**Proof of part (c)** Let  $\sigma_{ii} \in (0, 1)$ . We show that  $\tilde{\phi}_{N_i}^d(t) < \phi_{N_i}^d(t)$  implies  $\phi_i^{d^*}(\phi_i(t), \tilde{\phi}_{N_i}^d(t)) > \phi_i^{d^*}(\phi_i(t), \phi_{N_i}^d(t))$  by distinguishing between three cases.

- (i) Suppose  $\tilde{\phi}_{N_i}^d(t) \leq \phi_i(t) \leq \phi_{N_i}^d(t)$ , where at least one of the two inequalities is strict. Proposition 1 part (a) and part (b) imply that  $\phi_i^{d^*}(\phi_i(t), \tilde{\phi}_{N_i}^d(t)) \geq \phi_i(t) \geq \phi_i^{d^*}(\phi_i(t), \phi_{N_i}^d(t))$ , where strictness of each inequality carries over.
- (ii) Suppose  $\tilde{\phi}_{N_i}^d(t) < \phi_{N_i}^d(t) \leq \phi_i(t)$  and let  $x^*$  be the best response to  $\phi_{N_i}^d(t)$ , i.e. the solution to the maximization problem  $\max_{x \in \mathcal{I}} u_i(x) + f(x)$  as above. Now, consider  $\tilde{\phi}_{N_i}^d(t)$  instead of  $\phi_{N_i}^d(t)$  and define  $\tilde{f} : \mathbb{R} \mapsto \mathbb{R}$  s.t.  $\tilde{f}(x) = v_i(\sigma_{ii}x + (1 - \sigma_{ii})\tilde{\phi}_{N_i}^d(t))$ . Then the optimization problem (4) becomes (analogously to above),

$$\max_{x \in \mathcal{I}} u_i(x) + \tilde{f}(x). \quad (\text{B.2})$$

The optimal solution to (B.2) is denoted by  $\tilde{x}^*$ . Note that  $f$  and  $\tilde{f}$  are strictly concave for  $\sigma_{ii} \in (0, 1)$ . Since  $\tilde{\phi}_{N_i}^d(t) < \phi_{N_i}^d(t)$  we have that  $\tilde{f}(x)$  is a right-shifted version of  $f$ , i.e.  $\tilde{f}(x) = f(x - \delta)$  for  $\delta := \frac{1 - \sigma_{ii}}{\sigma_{ii}}(\phi_{N_i}^d(t) - \tilde{\phi}_{N_i}^d(t)) > 0$ . Thus, it holds that  $\tilde{f}'(x) > f'(x)$  for any  $x$  (since  $\tilde{f}'$  and  $f'$  are strictly decreasing).

For an interior solution  $x^*$ , which satisfies the first order condition  $f'(x^*) = -u_i'(x^*)$ , this implies  $\tilde{f}'(x^*) > -u_i'(x^*)$ . Hence,  $u_i'(x) + \tilde{f}'(x) > 0$ , i.e. the target function  $u_i(x) + \tilde{f}(x)$  is increasing at  $x^*$ , which yields  $\tilde{x}^* > x^*$  if  $x^*$  is interior. If  $x^*$  is not interior, we have  $x^* = \max \mathcal{I}$  (by Proposition 1 part (b)) and hence  $u_i'(x^*) + f'(x^*) \geq 0$ . Again,  $\tilde{f}'(x) > f'(x)$  implies  $u_i'(x^*) + \tilde{f}'(x^*) > 0$ , which leads to a boundary solution  $\tilde{x}^* = x^* = \max \mathcal{I}$ . However, this case is excluded by the assumption that one of the solutions must be interior. (This shows the claim in case ii since  $\tilde{x}^* = \phi_i^{d^*}(\phi_i(t), \tilde{\phi}_{N_i}^d(t)) > \phi_i^{d^*}(\phi_i(t), \phi_{N_i}^d(t)) = x^*$  if  $x^*$  is interior.)

- (iii) Suppose that  $\phi_i(t) \leq \tilde{\phi}_{N_i}^d(t) < \phi_{N_i}^d(t)$ . The arguments that establish that  $\phi_i^{d^*}(\phi_i(t), \tilde{\phi}_{N_i}^d(t)) > \phi_i^{d^*}(\phi_i(t), \phi_{N_i}^d(t))$  are fully analogous to Case (ii) above.

The three cases show the first direction (“IF”) of Proposition 1 part (c). To establish the other direction (“ONLY IF”) note first that  $\phi_{N_i}^d(t) < \tilde{\phi}_{N_i}^d(t)$  implies that  $\tilde{x}^* < x^*$  by the three cases above when exchanging  $\tilde{\phi}_{N_i}^d(t)$  and  $\phi_{N_i}^d(t)$ . Moreover,  $\tilde{\phi}_{N_i}^d(t) = \phi_{N_i}^d(t)$  implies  $\tilde{x}^* = x^*$  because there is a unique solution to the two coinciding maximization problems.  $\square$

## B.2 Proof of Proposition 2

From equation (2), it follows that  $\forall i \in N$ ,  $\phi_i(t+1)$  is linear in  $\phi_i^d(t)$ , thus  $v_i(\phi_i(t+1) | \phi_i(t))$  is concave in  $\phi_i^d(t)$  (by A3). This implies that the target functions of the optimization problems of all parents are concave (and continuous). Since also the displayed trait choice set is compact and convex, a non-empty, upper hemicontinuous and convex set of displayed trait best replies exists for any parent (Berger’s Theorem of the Maximum). Thus, a fixed point, i.e. a Nash equilibrium, exists (Kakutani’s Fixed Point Theorem).  $\square$

### B.3 Proof of Proposition 3

(a) That in any steady state, parents choose their adopted trait as displayed trait is directly implied by Proposition 1.

(b) By the definition of steady states and the trait formation rule (2), it follows that the set of steady states coincides with the set  $\{\Phi \in \mathcal{I}^n \mid \Sigma\Phi = \Phi\}$ . Hence, it is immediate that if the traits of all members of an essential communication class are identical, then  $\Sigma_L\Phi_L = \Phi_L$ , where  $\Sigma_L$  is the restriction of  $\Sigma$  to some essential communication class  $L$ , and  $\Phi_L$  is its vector of adopted traits restricted to that set. We proceed by showing that steady state traits cannot differ within an essential communication class. To show a contradiction, suppose that for an essential communication class  $L \in \mathcal{P}(\Sigma)$ ,  $|L| \geq 2$ , there exists  $i, j \in L$  with  $\phi_i \neq \phi_j$ . Denote by  $\bar{\phi}_L := \max\{\phi_i \mid i \in L\}$  the maximal trait in communication class  $L$ . Since  $L$  is a communication class, it follows that there exists an  $i \in \{l \in L \mid \phi_l = \bar{\phi}_L\}$  and a  $j \in \{l \in L \mid \phi_l \neq \bar{\phi}_L\}$  such that  $\sigma_{ij} > 0$ . Moreover, due to maximality of  $\bar{\phi}_L$  and the fact that  $L$  is essential,  $\sigma_{ik} = 0$  for all  $k \in N$  with  $\phi_k > \bar{\phi}_L$ . Thus,  $e'_i \Sigma \Phi_L \neq \phi_i$  implying that this cannot be a steady state ( $e_i$  denotes the  $i$ -th unit vector).

(c) Suppose that for some inessential communication class  $L' \in \mathcal{P}(\Sigma)$  with connections to other dynasties  $J := \{j \in N \mid i \rightarrow j, i \in L'\}$  the set of traits  $\Phi_{L'}$  is not included in  $\text{conv}(\phi_j \mid j \in J)$ . W.l.o.g. we have  $\bar{\phi}_{L'} := \max\{\phi_i \mid i \in L'\} > \max\{\phi_j \mid j \in J\}$ . Since  $L'$  is a communication class and is inessential with all outside connections being to dynasties with traits strictly less than  $\bar{\phi}_{L'}$ , we get (similarly to (b)) for some player  $k \in \{i \in L' \mid \phi_i = \bar{\phi}_{L'}\}$  that there exists  $j \in N$  and  $\phi_j < \bar{\phi}_{L'}$  such that  $\sigma_{kj} > 0$ . Again, due to maximality of  $\bar{\phi}_{L'}$  and all other connections being to dynasties with traits strictly less than  $\bar{\phi}_{L'}$ , we get that  $e'_k \Sigma \Phi_{L'} \neq \phi_k$ , implying that this cannot be a steady state. Hence, all traits of the dynasties in inessential communication classes  $L' \in \mathcal{P}(\Sigma)$  are convex combinations of the traits of the communication classes  $L \in \mathcal{P}(\Sigma)$  such that  $J \cap L \neq \emptyset$ .

### B.4 Proof of Proposition 4

To prove the proposition, we will apply the following Lemma (see e.g. Friedberg and Insel, 1992).

**Lemma B1** (Convergence). *Let  $A$  be a square matrix with complex or real entries. Then, the sequence  $\{A^t\}_{t \rightarrow \infty}$  converges if and only if the following two conditions are satisfied.*

- (i) *If  $\lambda$  is an eigenvalue of  $A$ , then either  $\lambda = 1$  or  $\lambda$  lies in the open unit disc of the complex plane, i.e.  $|\lambda| \in (-1, 1)$ .*
- (ii) *If 1 is an eigenvalue of  $A$ , then its algebraic multiplicity equals its geometric multiplicity.*

Let us denote by  $\Lambda(A)$  the set of eigenvalues of a matrix  $A$  and let  $\lambda(A) \in \Lambda(A)$ . Moreover, if  $z$  is a complex number, then we denote by  $\text{Re}(z)$  the real part and by  $\text{Im}(z)$  the imaginary part of  $z$ .

**Proof of part (a)** We show that condition (i) of Lemma B1 is satisfied for symmetric positive definite matrices  $\Sigma$ . To see this, note first that by definition  $M = \Sigma(I + B\Sigma)^{-1}(I + B) = (B + \Sigma^{-1})^{-1}(I + B)$ ,<sup>24</sup> which implies that  $M$  is invertible and  $M^{-1} = (I + B)^{-1}(B + \Sigma^{-1})$ . Let  $\tilde{B}$  denote the diagonal matrix  $\tilde{B} := (I + B)^{-1}$ . Thus, for all  $i = 1, \dots, n$  the entries of  $\tilde{B}$  given by  $\tilde{b}_{ii} = \frac{1}{1 + \sigma_{ii}\beta_i}$  are such that  $\tilde{b}_{ii} \in (0, 1)$ . Then,  $\tilde{B}B = I - \tilde{B}$ , and

$$M^{-1} = \tilde{B}(B + \Sigma^{-1}) = I - \tilde{B} + \tilde{B}\Sigma^{-1} = I + \tilde{B}(\Sigma^{-1} - I). \quad (\text{B.3})$$

First, note that since  $\Sigma$  is assumed to be symmetric positive definite, so is  $\Sigma^{-1}$  and  $(\Sigma^{-1} - I)$  (see below) and the eigenvalues of these matrices are real and positive.

Second, the matrices  $\tilde{B}(\Sigma^{-1} - I) = \tilde{B}^{1/2}[\tilde{B}^{1/2}(\Sigma^{-1} - I)]$  and  $\tilde{B}^{1/2}(\Sigma^{-1} - I)\tilde{B}^{1/2}$  have the same eigenvalues,<sup>25</sup> where  $\tilde{B}^{1/2}$  is the diagonal matrix with entries  $(\tilde{B}^{1/2})_{ii} = \sqrt{\tilde{b}_{ii}}$ . Moreover, it is easily checked that  $\tilde{B}^{1/2}(\Sigma^{-1} - I)\tilde{B}^{1/2}$  is positive definite and symmetric, i.e. has only positive real eigenvalues. Thus, also the eigenvalues of  $M^{-1}$  (and hence those of  $M$ ) are real and positive.

Now, since  $\Sigma$  is row stochastic, we have  $|\lambda(\Sigma)| \leq 1$ , which implies that  $\lambda(\Sigma^{-1}) \geq 1$ . Thus,  $\lambda(\Sigma^{-1} - I) \geq 0$  (subtraction of  $I$  decreases all eigenvalues by 1). By above, we have  $\lambda(\tilde{B}(\Sigma^{-1} - I)) \geq 0$ , which implies  $\lambda(I + \tilde{B}(\Sigma^{-1} - I)) \geq 1$ , i.e.  $\lambda(M^{-1}) \geq 1$ , and hence all eigenvalues of  $M$  are real and located in the interval  $(0, 1]$ . Furthermore, since  $M$  has row sum one (see Lemma C2, using  $x = (1, 1, \dots, 1)'$ ), at least one eigenvalue must be equal to 1. Note that, the number of 1-eigenvalues as well as the number of associated linear independent eigenvectors equals the number of essential communication classes of  $M$ . Hence condition (ii) of Lemma B1 is also satisfied. Thus,  $M^t$  converges, i.e.  $M^\infty := \lim_{t \rightarrow \infty} M^t$  exists, and since 1 is an eigenvalue of  $M$ ,  $M^\infty \neq \mathbf{0}$ . Denoting  $\Phi(\infty) := M^\infty\Phi(0)$  it is easy to see that  $\Phi(\infty)$  is a steady state since  $M\Phi(\infty) = MM^\infty\Phi(0) = M^\infty\Phi(0) = \Phi(\infty)$ .

**Proof of part (b)** Assume that the condition in the statement is not satisfied, i.e. let there be an eigenvalue  $\tilde{\lambda}(\Sigma)$  that satisfies  $\text{Re}(\tilde{\lambda}(\Sigma)) < |\tilde{\lambda}(\Sigma)|^2$ . The latter is equivalent to  $\text{Re}(\tilde{\lambda}^{-1}(\Sigma)) < 1$ , simply because  $z^{-1} = \frac{\text{Re}(z)}{\text{Re}^2(z) + \text{Im}^2(z)} + \frac{-\text{Im}(z)}{\text{Re}^2(z) + \text{Im}^2(z)}i$  and  $|z|^2 = \text{Re}^2(z) + \text{Im}^2(z)$  for any complex number  $z \in \mathbb{C}$ . By assumption,  $\Sigma$  is invertible and we note that  $\tilde{\lambda}^{-1}(\Sigma)$  is an eigenvalue of  $\Sigma^{-1}$ . Now, let for each  $i$ ,  $\beta_i \equiv \frac{k}{\sigma_{ii}}$ ,  $k \in \mathbb{R}$ , so that  $B = kI$ . We show that if  $k$  is large enough, then  $M$  has an eigenvalue with absolute value larger than 1 and hence condition (i) of Lemma B1 is violated.

To do so, we will use  $M^{-1} = (I + B)^{-1}(B + \Sigma^{-1}) = (I + kI)^{-1}(kI + \Sigma^{-1}) = ((1 + k)I)^{-1}(kI + \Sigma^{-1}) = \frac{1}{1+k}(kI + \Sigma^{-1})$ . Now, since  $\text{Re}(\tilde{\lambda}(\Sigma^{-1})) = \text{Re}(\tilde{\lambda}^{-1}(\Sigma)) < 1$ , we have  $\text{Re}(\tilde{\lambda}(kI + \Sigma^{-1})) < 1 + k$ , because  $\tilde{\lambda}(kI + \Sigma^{-1}) = k + \tilde{\lambda}(\Sigma^{-1})$ . For  $k$  large enough, we must have  $|\tilde{\lambda}(kI + \Sigma^{-1})| < 1 + k$ .<sup>26</sup> To see that this must hold, denote

<sup>24</sup>That this representation is well defined if  $\Sigma$  is positive definite has been discussed in footnote 18.

<sup>25</sup>This holds since for any two  $n \times n$  matrices  $A, B$  the eigenvalues of  $AB$  are the same as the eigenvalues of  $BA$ , although the eigenvectors may differ.

<sup>26</sup>If  $\tilde{\lambda}^{-1}(\Sigma)$  is a real number, then this holds trivially.

$\epsilon := 1 - \operatorname{Re}(\tilde{\lambda}(\Sigma^{-1}))$  and we get:

$$\begin{aligned} |\tilde{\lambda}(kI + \Sigma^{-1})|^2 &= \operatorname{Re}^2(\tilde{\lambda}(kI + \Sigma^{-1})) + \operatorname{Im}^2(\tilde{\lambda}(kI + \Sigma^{-1})) \\ &= (1 - \epsilon + k)^2 + \operatorname{Im}^2(\tilde{\lambda}(\Sigma^{-1})) \\ &= (k + 1)^2 + \operatorname{Im}^2(\tilde{\lambda}(\Sigma^{-1})) + \epsilon^2 - 2\epsilon - 2\epsilon k, \end{aligned}$$

which is smaller than  $(1 + k)^2$  for  $k > \frac{\operatorname{Im}^2(\tilde{\lambda}(\Sigma^{-1})) + \epsilon^2 - 2\epsilon}{2\epsilon}$ . Thus, we get for  $k$  large enough,

$$\frac{1}{1+k} |\tilde{\lambda}(kI + \Sigma^{-1})| = \left| \tilde{\lambda} \left( \frac{1}{1+k} (kI + \Sigma^{-1}) \right) \right| = |\tilde{\lambda}(M^{-1})| < 1$$

and hence  $|\tilde{\lambda}(M)| > 1$  so that condition (i) of Lemma B1 is violated.  $\square$

## B.5 Proof of Proposition 5

As by Lemma B1 above, for the convergence of the powers of a matrix  $A$  it is sufficient that 1 is exactly one eigenvalue of  $A$  and all other eigenvalues are in the interval  $(-1, 1)$ . To prove the proposition, we will in a first step apply the Perron-Frobenius Theorem (henceforth: PFT) for a regular row-stochastic matrix  $A$ : (i) The spectral radius (the largest eigenvalue in absolute value) of  $A$  is 1. (ii) For all other eigenvalues  $\lambda$  it holds that  $|\lambda| < 1$ . (iii) The eigenvalue 1 is simple. Consider any row stochastic  $\Sigma$  such that  $\Sigma$  is irreducible with strictly positive diagonal. This implies that  $\Sigma$  is regular, so that by the PFT for regular row stochastic matrices,  $\Sigma$  has simple eigenvalue 1 and all other eigenvalues are in  $(-1, 1)$ .

Let us now consider the transformations  $M = \Sigma(I + B\Sigma)^{-1}(I + B)$ . In a first step, we have to guarantee that  $I + B\Sigma$  is invertible, so that  $M$  exists. Note that strict diagonal dominance would be sufficient for non-singularity. For strict diagonal dominance, we require that  $1 + \beta_i \left( \sigma_{ii} - \sum_{j \in N_i} \sigma_{ij} \right) > 0$  holds for every  $i \in N$ . Since  $\Sigma$  has a strictly positive diagonal, this is always satisfied if e.g.  $\beta \leq \mathbf{1}$ .

Given this, it follows again by the continuity of the eigenvalues that there exists a non-empty neighborhood  $N(\mathbf{0} | \Sigma) \subset \mathbb{R}_+^n$  such that  $\forall \beta \in N(\mathbf{0} | \Sigma) \cup \mathbf{0}$  both  $I + B\Sigma$  is strictly diagonally dominant and  $M$  has exactly one eigenvalue equal 1 and  $n - 1$  eigenvalues in the interval  $(-1, 1)$ . Thus,  $M^t$  converges.  $\square$

## B.6 Proof of Proposition 6

For symmetric positive definite and row stochastic matrices  $\Sigma$ , convergence of  $\Sigma^t$  for  $t \rightarrow \infty$  is trivially implied and convergence of  $M^t$  for  $t \rightarrow \infty$  is already established by Proposition 4. To show that convergence of  $M^t$  is slower than convergence of  $\Sigma^t$  for  $t \rightarrow \infty$ , we show that all eigenvalues of  $M$  are real and  $\lambda_k(\Sigma) < \lambda_k(M)$  for all  $2 \leq k \leq K$ .

By (B.3) we have that  $M^{-1} = I + \tilde{B}(\Sigma^{-1} - I)$  with  $\tilde{B}$  being a diagonal matrix with entries  $0 < \tilde{b}_{ii} = \frac{1}{1 + \sigma_{ii}\beta_i} < 1$  for all  $\beta_i > 0$ . As in the proof of Proposition 4 we have,

$$\lambda_k(M^{-1}) = 1 + \lambda_k(\tilde{B}(\Sigma^{-1} - I)) = 1 + \lambda_k(B^{1/2}(\Sigma^{-1} - I)B^{1/2}).$$

Moreover, by the proof of Proposition 4,  $(\Sigma^{-1} - I)$  and  $\tilde{B}^{1/2}(\Sigma^{-1} - I)\tilde{B}^{1/2}$  are symmetric and positive definite.

Since  $(\tilde{B}^{1/2})^* = \tilde{B}^{1/2}$  and  $\tilde{B}^{1/2}$  is non-singular,<sup>27</sup> and  $(\Sigma^{-1} - I)$  is symmetric, we get by Theorem 1 in Ostrowski (1959) (see also Horn and Johnson, 2010) that  $\lambda_k(\tilde{B}^{1/2}(\Sigma^{-1} - I)\tilde{B}^{1/2}) = \theta_k \lambda_k(\Sigma^{-1} - I)$ , where  $\theta_k$  are real numbers such that  $\lambda_K(\tilde{B}^{1/2}\tilde{B}^{1/2}) \leq \theta_k \leq \lambda_1(\tilde{B}^{1/2}\tilde{B}^{1/2})$ . Since  $\tilde{B}^{1/2}\tilde{B}^{1/2} = \tilde{B}$  is diagonal with entries  $0 < \tilde{b}_{ii} < 1$ , it holds that  $\lambda_k(\tilde{B}^{1/2}(\Sigma^{-1} - I)\tilde{B}^{1/2}) < \lambda_k(\Sigma^{-1} - I)$  for all  $k$  such that  $\lambda_k(\Sigma^{-1} - I) > 0$ . The latter is satisfied for  $\lambda_k(\Sigma) < 1$ , and thus for all  $\lambda_k(\Sigma)$  such that  $2 \leq k \leq K$ .

Hence, for  $2 \leq k \leq K$  :

$$\begin{aligned} \lambda_k(M^{-1}) &= 1 + \lambda_k(\tilde{B}(\Sigma^{-1} - I)) \\ &< 1 + \lambda_k(\Sigma^{-1} - I) = 1 + \lambda_k(\Sigma^{-1}) - 1 = \lambda_k(\Sigma^{-1}), \end{aligned}$$

which implies that  $\lambda_k(M) > \lambda_k(\Sigma)$  for all  $2 \leq k \leq K$  proving the statement and implying that convergence of  $M^t$  is slower than convergence of  $\Sigma^t$  for  $t \rightarrow \infty$ .

To show the second part, consider the limit when all socialization incentives grow,  $\beta_i \rightarrow \infty$  for all  $i = 1, \dots, n$ . Then,  $\tilde{b}_{ii} = \frac{1}{1 + \sigma_{ii}\beta_i} \rightarrow 0$ , and, hence,  $M^{-1} = I + \tilde{B}(\Sigma^{-1} - I) \rightarrow I$ . Thus,  $\lambda_k(M) \rightarrow 1$  for all eigenvalues  $\lambda_k$  of  $M$ ,  $k = 1, \dots, K$ .

## C Convergence in the General Case

We show in this part of the appendix how convergence conditions for more general utility functions can be obtained. In particular, the convergence conditions found for quadratic utility (presented in Section 4.2) generalize nicely to utility functions that satisfy Assumptions A1-A3. In this case, best reply displayed cultural traits  $\Phi^{d^*}(t)$  need not be linear in the true cultural traits  $\Phi(t)$ . Since we assume that in each period a Nash equilibrium is played, the dynamics are non-linear, too. In particular, the law of motion of the dynamics depends on the vector of cultural traits  $\Phi(t)$ . First, we show how such dynamics can be represented by a linear system, which is analogous but more general than the process  $\Phi(t+1) = M\Phi(t)$  that we derived in the quadratic utility case. Then we will derive a condition on the network structure  $\Sigma$  which is sufficient for convergence and finally we prove this result.

### C.1 Linearizing the Dynamical System

In this section we show how the dynamics of cultural traits can be presented by a linear system, a product of *time-dependent* matrices  $M(t)$ . For this purpose, we define a map  $B^*$  which picks one Nash equilibrium each period.

**Corollary C1** (Nash Equilibrium Map). *There exists a Nash equilibrium map  $B^* : \mathcal{I}^n \mapsto \mathbb{R}_+^n$ , such that for every  $i \in N$  and for every  $t \in \mathbb{N}$ ,  $B^*(\Phi(t)) = (b_1^*(t), \dots, b_n^*(t))'$  satisfies*

$$\phi_i^{d^*}(t) - \phi_i(t) = b_i^*(t) \cdot (\phi_i(t) - \phi_i^*(t+1))$$

where  $\phi_i^*(t+1) := \sum_{j \in N} \sigma_{ij} \phi_j^{d^*}(t)$ . This map has the property that  $b_i^*(t) \geq 0$  for all  $t \in \mathbb{N}$ . Moreover, if for all  $i \in N$ ,  $\sigma_{ii} = 0$ , then  $b_i^*(t) = 0$ ,  $\forall t \in \mathbb{N}$ .

<sup>27</sup>The asterisk denotes the complex conjugate transpose.

*Proof.* Follows immediately from the best reply characterization of Proposition 1 and the Nash equilibrium existence of Proposition 2.  $\square$

The Nash equilibrium map simply represents the Nash equilibrium displayed trait choices in terms of their deviations from the adults' traits relative to the deviation of the children's adopted traits from the socialization targets. This representation can equivalently be written as  $\phi_i^{d^*}(t) + b_i^*(t)\Sigma_i\Phi^{d^*}(t) = (1 + b_i^*(t))\phi_i(t)$ , for every  $i \in N$ . Defining  $B(t) := \text{diag}(b_1^*(t), \dots, b_n^*(t))$ , we thus obtain,

$$(I + B(t)\Sigma)\Phi^{d^*}(t) = (I + B(t))\Phi^*(t)$$

so that

$$\Phi^{d^*}(t) = (I + B(t)\Sigma)^{-1}(I + B(t))\Phi^*(t)$$

and hence

$$\Phi^*(t+1) = \Sigma(I + B(t)\Sigma)^{-1}(I + B(t))\Phi^*(t).$$

For this representation to be well-defined, it is sufficient that either  $\Sigma$  is diagonally dominant (since then  $I + B(t)\Sigma$  is then strictly diagonally dominant, thus invertible) or symmetric positive semidefinite (the assumptions used in our result below will imply in particular that  $\Sigma$  is symmetric positive definite).

Finally, denoting  $M(t) := \Sigma(I + B(t)\Sigma)^{-1}(I + B(t))$ , it follows that

$$\Phi^*(t+1) = M(t) \dots M(0)\Phi(0) = M(t,0)\Phi(0), \quad t \in \mathbb{N} \setminus \{0\} \quad (\text{C.1})$$

where  $M(t,0)$  denotes the *backward accumulation*  $M(t,0) := M(t) \cdot M(t-1) \cdot \dots \cdot M(0)$ .

## C.2 Convergence Result

The representation of the dynamics derived in Appendix C.1 allows us to resort to linear algebra results on the convergence of left products of matrices. Specifically, Lorenz (2005, 2006) provides convergence results for left products of row stochastic matrices—while as (for our specific context) not sufficient results are available on the left product convergence of more general matrices (that have row sum one, but with possibly negative entries). However, to guarantee that the individual matrices  $M(t)$  are row stochastic in every period  $t \in \mathbb{N}$ , we have to endow the social learning matrix  $\Sigma$  with sufficient structure, which is given by the following definition.

**Definition 1** (Symmetric Ultrametric Matrix). *A  $n \times n$ -matrix  $\Sigma$  is symmetric ultrametric if*

- (i)  $\Sigma$  is symmetric,
- (ii)  $\sigma_{ii} \geq \max\{\sigma_{ij} : j \in N_i\}, \forall i \in N$ ,
- (iii)  $\sigma_{ij} \geq \min\{\sigma_{ik}; \sigma_{kj}\}, \forall i, j, k \in N$ .

Property (i) means that interaction weights are symmetric, i.e.  $\forall i, j \in N \sigma_{ij} = \sigma_{ji}$ . Property (ii) means that among all adults, the parents have the largest socialization influence on their children. In general, the third property requires a sort of consistency of the socialization patterns. It states that for any triple  $i, j, k \in N$ , if the socialization

influence of  $j$  on child  $i$  is strictly smaller than that of  $k$  on child  $i$ , then it must not hold that  $k$  has a strictly larger socialization influence on child  $j$  than on child  $i$  (since  $\sigma_{kj} = \sigma_{jk}$ ). This requirement can be interpreted as ruling out the existence of dynasties that have a ‘too dominant’ social learning influence on other dynasties.<sup>28</sup>

We now get the following convergence result.

**Proposition C1** (Convergence General). *If  $\Sigma$  is symmetric ultrametric, then the cultural traits in the society  $\Phi(t)$  converge to a steady state.*

*Proof.* The proof is presented in Subsection C.3 below.

In the proof of Proposition C1 we show that given that the social network  $\Sigma$  is symmetric ultrametric, the submatrix of each essential communication class will converge to a matrix of rank one (i.e. a consensus matrix). This is done by making use of the special structure of  $\Sigma$  to show that each element of the law of motion  $M(t)$  is row stochastic.<sup>29</sup> Second, we can show that the entries of  $M(t)$  corresponding to strictly positive entries of  $\Sigma$  can be bounded away from zero. This is due to the linearity of the determinants of the minors of  $M(t)$  in every  $b_i^*(t)$ , and the boundedness of  $b_i^*(t)$ . In the last step, we construct a sequence of sub-accumulations of  $M(t_{s+1}, t_s)_{s \in \mathbb{N}}$  such that for each element the minimal strictly positive entry can be uniformly bounded away from zero, which also implies “type-symmetry” and a strictly positive diagonal. Thus, we can then apply the convergence result by Lorenz (2005), which implies that the traits of each connected subset converge to the same point, i.e. all dynasties in a communication class reach a homogeneous state. Note that due to the assumption of symmetric ultrametric  $\Sigma$  all communication classes are essential and there is, hence, no rest of the world.

Thus, endowing the social learning matrix  $\Sigma$  with sufficient structure, we arrive at a general result: In the long-run the dynasties forming a communication class (i.e. forming a component of the social network) will reach a steady state, and will end up with the same cultural trait (cf. also Proposition 3). The traits across communication classes may differ. Thus, symmetric ultrametric and connected societies (i.e. where every dynasty is at least indirectly influenced by any other dynasty) will always converge to a melting pot society even if objective functions are fairly general. Finally, we note that the typical assumption in the literature is that oblique socialization is homogeneous in the sense that all members are weighted equally, which translates to  $\sigma_{ij} = \sigma_{ik}$  for all  $j, k \neq i$ . It is straightforward to see that this assumption along with symmetry implies that the interaction structure is symmetric ultrametric and hence by Proposition C1 convergence also obtains in these cases.

In the quadratic utility case studied in Section 4.2, we needed the social network  $\Sigma$  to be symmetric positive definite to ensure convergence, which is satisfied in a diagonally dominant network. When utility is more general, we have shown here a qualitatively similar result: a sufficient condition on the network for convergence is that it is symmetric ultrametric, which also means that the social network is symmetric, that the

---

<sup>28</sup>The third property (iii) of symmetric ultrametric matrices is a strong assumption, which we do not need when studying the case of quadratic utility (cf. Proposition 4).

<sup>29</sup>For literature on inverses of symmetric ultrametric matrices see Nabben and Varga (1993, 1994), Martinez et al. (1994), and for results on inverse-positive matrices see e.g. Fujimoto and Ranade (2004).

diagonal is sufficiently large, and, additionally, that no dynasty has dominant influence on others.

The assumption of symmetric ultrametric networks guarantees that the law of motion  $M(t)$  is row stochastic for all  $t \in \mathbb{N}$ . However, the necessity to guarantee that all  $M(t)$  are row stochastic significantly reduces the convergence path types that we can analytically address. Row stochasticity of  $M(t)$  implies that dynamics of cultural traits are such that all next-period cultural traits lie in the interval formed by the minimum and the maximum trait intensity of the current period. Thus, we can state the following straightforward Corollary.

**Corollary C2.** *If  $\Sigma$  is symmetric ultrametric, then  $\phi_i(t+1) \in \text{conv}\{\phi_j(t), j \in N\}$  for all  $i \in N$  and for all  $t \in \mathbb{N}$ .*

*Proof.* Corollary C2 follows directly from Proposition C1 since in the proof of which it is shown that  $M(t)$  is row stochastic for each  $t \in \mathbb{N}$  and hence  $\phi_i(t+1) = e_i' M \Phi(t) \in \text{conv}\{\phi_j(t), j \in N\}$  for all  $i \in N$ , where  $e_i$  is the  $i$ -th unit vector.  $\square$

Corollary C2 particularly implies that each dynasty will converge to a cultural trait that is a mixture of the initial cultural traits in the sense that it lies in the convex hull of the cultural traits in period 0.

### C.3 Proof of Proposition C1

This proof is organized into three essential steps. In the first step, we will show that if  $\Sigma$  is symmetric ultrametric, then  $M(t)$  is row stochastic for every  $t \in \mathbb{N}$ . In the second step we will show that for every  $i, j \in N$  with  $\Sigma_{ij} > 0$ , there exists a  $\delta_{ij} > 0$  such that for every  $t \in \mathbb{N}$ ,  $m_{ij}(t) \geq \delta_{ij}$ . We use these results to show in the third step that the backward accumulation matrices are type-symmetric and have a strictly positive diagonal. This allows us to apply Theorem 2 of Lorenz (2005) to conclude that the desired convergence result holds. For the first step, we also need the following.

**Lemma C2 (Unit Eigenvectors).** *Let  $\Sigma$  be symmetric positive definite. Then,  $\forall x \in \mathbb{R}^n$ ,  $\forall t \in \mathbb{N}$ ,  $M(t)x = x$  if and only if  $\Sigma x = x$  (i.e.  $x$  is a unit-eigenvector of  $M(t)$  if and only if  $x$  is a unit-eigenvector of  $\Sigma$ ).*

*Proof.* Note that  $M(t) = \Sigma(I + B(t)\Sigma)^{-1}(I + B(t)) = (\Sigma^{-1} + B(t))^{-1}(I + B(t))$ . That the latter representation is well-defined if  $\Sigma$  is positive definite follows since  $\Sigma$  is then invertible and also its inverse is positive definite. Thus, also  $\Sigma^{-1} + B(t)$  is positive definite and invertible. Given this, both the ‘if’ and the ‘only if’ direction of the proof can be directly seen from the following sequence of transformations:  $\Sigma x = x \Leftrightarrow x = \Sigma^{-1}x \Leftrightarrow (B(t) + I)x = (B(t) + \Sigma^{-1})x \Leftrightarrow M(t)x = (B(t) + \Sigma^{-1})^{-1}(B(t) + I)x = x$ .  $\square$

1. In the first step of the (main) proof, we show that if  $\Sigma$  is symmetric ultrametric, then  $M(t)$  is row stochastic for every  $t \in \mathbb{N}$ . To do so, note first that since  $\Sigma$  is symmetric ultrametric, it is also positive definite (see below). Hence, by Lemma C2 (and setting  $x = (1, 1, \dots, 1)'$ ) the row entries of  $M(t) = [m_{ij}(t)]$  sum up to one since the same holds for  $\Sigma$ . Thus,  $M(t)$  is row stochastic if and only if  $M(t)$  has non-negative entries (that is  $M(t) \geq 0$ ). Now, since  $I + B(t)$  is a diagonal matrix with strictly positive

entries (since  $B(t)$  is non-negative by Corollary C1),  $M(t) = \Sigma (I + B(t)\Sigma)^{-1} (I + B(t))$  is non-negative if and only if

$$\Sigma (I + B(t)\Sigma)^{-1} = (\Sigma^{-1} + B(t))^{-1}$$

is non-negative (that this representation is well defined if  $\Sigma$  is positive definite has been discussed in the proof of Lemma C2). In other words, we have to check whether  $\Sigma^{-1} + B(t)$  is inverse-positive.

Now, since  $\Sigma$  is symmetric ultrametric, it follows that its inverse is a diagonally dominant Stieltjes matrix (see Nabben and Varga, 1993, 1994 and Martinez et al., 1994), i.e. a real symmetric positive definite matrix with positive diagonal and negative off-diagonal entries. Thus, also  $\Sigma^{-1} + B(t)$  is a diagonally dominant Stieltjes matrix. In particular, it is an M-matrix, the class of which is inverse-positive (on this issue, see e.g. Fujimoto and Ranade, 2004). Hence,  $M(t)$  has only non-negative entries.

2. For the second step, we show first, that the map  $b_i^*(t) = b_i^*(\phi_i(t), \phi_{N_i}^d(t))$  is bounded for every  $i \in N$ .

**Lemma C3** (Boundedness of  $B^*(t)$ ). *Let Assumptions 1–3 hold. Then,  $\forall i \in N$   $b_i^*$  is bounded for every  $\phi_{N_i}^d(t), \phi(t) \in \mathcal{I}$ . In particular,*

$$\lim_{\phi_{N_i}^d(t) \rightarrow \phi_i(t)} b_i^*(\phi_i(t), \phi_{N_i}^d(t)) = \frac{\sigma_{ii} v_i''(\phi_i(t) | \phi_i(t))}{u_i''(\phi_i(t) | \phi_i(t)) + \sigma_{ii}^2 v_i''(\phi_i(t) | \phi_i(t))} < \infty.$$

*Proof.* Note that for  $x := \phi_i(t)$ ,  $y := \phi_{N_i}^d(t)$ , and  $f(x, y) := \phi_i^{d*}(x, y)$ ,  $b_i^*$  is defined by (see Corollary C1)

$$f(x, y) - x = b_i^*(x, y) ((1 - \sigma_{ii})x - (1 - \sigma_{ii})y). \quad (\text{C.2})$$

Let  $x \in \mathcal{I}$  be given and without loss of generality assume that  $y \geq x$ . First, note that for every  $y \in \mathcal{I}$  such that  $x \neq y$  it holds by Proposition 1 that  $0 \leq b_i^*(x, y) \leq \frac{1}{1 - \sigma_{ii}} \frac{x - x_{\min}}{x - y}$  for  $x_{\min} := \min\{z \in \mathcal{I}\}$ , since by Proposition 1  $x_{\min} \leq f(x, y) \leq x$ . Further, by Proposition 1 we get for  $\sigma_{ii} = 1$  that  $b_i^*(x, y) = 0$  for all  $y > x$ .

Hence we are left to show that  $\lim_{y \downarrow x} b_i^*(x, y) < \infty$  for  $\sigma_{ii} < 1$ . Since  $x$  is fixed, we denote  $f(y) := f(x, y)$ , abusing notation. We get from (C.2),

$$\lim_{y \downarrow x} b_i^*(x, y) = \lim_{y \downarrow x} \frac{1}{1 - \sigma_{ii}} \frac{f(y) - x}{y - x} = -\frac{1}{1 - \sigma_{ii}} f'(x),$$

given differentiability of  $f$  at the point  $x$ , which we show subsequently. By the first order condition,  $f(y)$  solves  $u_i'(f(y) | x) + \sigma_{ii} v_i'(\sigma_{ii}(f(y) + (1 - \sigma_{ii})y) | x) = 0$ . With the implicit function theorem,

$$f'(x) = -\frac{(1 - \sigma_{ii})\sigma_{ii} v_i''(\sigma_{ii}(f(x) + (1 - \sigma_{ii})x) | x)}{u_i''(f(x) | x) + \sigma_{ii}^2 v_i''(\sigma_{ii}(f(x) + (1 - \sigma_{ii})x) | x)}. \quad (\text{C.3})$$

By Proposition 1, we have  $f(x) = x$ , and hence by Assumption 1(c) the right-hand side is well defined implying differentiability of  $f$  at  $x$ . We get

$$\begin{aligned} \lim_{y \downarrow x} b_i^*(x, y) &= -\frac{1}{1 - \sigma_{ii}} f'(x) \\ &= \frac{\sigma_{ii} v_i''(x|x)}{u_i''(x|x) + \sigma_{ii}^2 v_i''(x|x)}, \end{aligned}$$

which is by A3 positive and bounded. □

Now, we continue to show that for every  $i, j \in N$  with  $\sigma_{ij} > 0$  there exists a  $\delta_{ij} > 0$  such that  $m_{ij}(t) \geq \delta_{ij}$  for every  $t \in \mathbb{N}$ . Again, since  $I + B(t)$  is a diagonal matrix with strictly positive entries, we can restrict our attention to the matrix  $(\Sigma^{-1} + B(t))^{-1} =: A(t) = [a_{ij}(t)]$ . Consider any  $i, j \in N$  such that  $\sigma_{ij} > 0$ . Since  $A(t)$  is non-negative by step (1), it follows that  $\text{sign}(a_{ij}(t)) \in \{0, \text{sign}(\sigma_{ij})\}$ .

Let us rule out the case  $\text{sign}(a_{ij}(t)) = 0$  for  $\sigma_{ij}(t) > 0$ . To do so, let us compare

$$a_{ij}(t) = (-1)^{i+j} \frac{|\Sigma^{-1} + B(t)|_{ji}}{|\Sigma^{-1} + B(t)|} \quad \text{vs.} \quad (-1)^{i+j} \frac{|\Sigma^{-1}|_{ji}}{|\Sigma^{-1}|} = \sigma_{ij} \quad (\text{C.4})$$

where  $|\cdot|_{ji}$  denotes the determinant of the  $(n-1) \times (n-1)$  matrix derived from an  $n \times n$  matrix such that the  $j$ -th row and the  $i$ -th column are deleted. Note that since  $\Sigma$  is positive definite, the same holds for its inverse and  $\Sigma^{-1} + B(t)$ . It follows that the determinants of the matrices  $\Sigma^{-1}$  and  $\Sigma^{-1} + B(t)$  are strictly positive and hence the denominators of (C.4) are strictly positive.

Moreover, we have that for all  $i, j \in N$ ,  $|\Sigma^{-1} + \text{diag}(b_1^*(t), \dots, b_n^*(t))|_{ji}$  and  $|\Sigma^{-1} + \text{diag}(b_1^*(t), \dots, b_n^*(t))|$  are linear in every individual element of  $\{b_1^*(t), \dots, b_n^*(t)\}$  (to verify this most easily, consider the Leibniz formula). Since we have  $|\Sigma^{-1} + B(t)| \geq 0$  for all  $b_1^*(t), \dots, b_n^*(t) \geq 0$ , it holds that

$$\frac{\partial(-1)^{i+j} |\Sigma^{-1} + B(t)|_{ji}}{\partial b_k} \geq 0 \quad (\text{C.5})$$

and

$$\frac{\partial |\Sigma^{-1} + B(t)|}{\partial b_k} \geq 0, \quad (\text{C.6})$$

because otherwise the determinant would switch signs for  $b_k \rightarrow \infty$  due to linearity in  $b_k$ , for all  $k \in \{1, \dots, n\}$ .

Now, since  $b_i^*$  is bounded by Lemma C3, we have  $b_k(t) \in [0, \bar{b}]$  for all  $t \in \mathbb{N}$ . By linearity of  $|\Sigma^{-1} + B(t)|_{ij}$  and  $|\Sigma^{-1} + B(t)|$  in  $b_k(t)$  for all  $k \in \{1, \dots, n\}$  and compactness of  $[0, \bar{b}]$ , we thus get existence of a minimum:

$$\hat{\delta}_{ij} := \min_{k \in \{1, \dots, n\}} \min_{b_k(t) \in [0, \bar{b}]} = (-1)^{i+j} \frac{|\Sigma^{-1} + B(t)|_{ji}}{|\Sigma^{-1} + B(t)|} \leq (-1)^{i+j} \frac{|\Sigma^{-1}|_{ji}}{|\Sigma^{-1}|}$$

Moreover,  $0 < \hat{\delta}_{ij}$  since both of nominator and denominator are bounded and strictly positive due to (C.5) and (C.6) and because of boundedness of  $b_k$ .

Thus, if  $\sigma_{ij} > 0$ , then  $a_{ij}(t) \geq \hat{\delta}_{ij}$  for all  $t \in \mathbb{N}$ . Multiplication with the diagonal matrix  $I + B(t)$  does not change this fact, even though the minimum might be attained at different values of  $b_k \in [0, \bar{b}]$  and  $k \in \{1, \dots, n\}$ . Thus, for all  $i, j \in \mathbb{N}$  such that  $\sigma_{ij} > 0$ , there exists a  $\delta_{ij} > 0$  such that  $m_{ij}(t) \geq \delta_{ij}$  for all  $t \in \mathbb{N}$ .

3. In the last step, we show that given the above, the left product of the matrices  $M(t)M(t-1)\dots M(0)$  converges such that the adopted traits of all dynasties of a connected subset are identical (respectively, the communication classes in  $\mathcal{P}(\Sigma)$  reach a consensus). Recall that  $M(t', t)$  denotes the accumulation  $M(t', t) = M(t')M(t'-1)\dots M(t)$  and  $P_\Sigma(i) \subseteq N$  denotes the element of the partition  $\mathcal{P}(\Sigma)$  which  $i$  belongs to, i.e.  $P_\Sigma(i)$  is such that  $P_\Sigma(i) \in \mathcal{P}(\Sigma)$  and  $i \in P_\Sigma(i)$ .

First, note that all communication classes of  $\Sigma$  are essential by symmetry of  $\Sigma$ . By the definition of  $\mathcal{P}(\Sigma)$ , we have that for all  $L \in \mathcal{P}(\Sigma)$  and for all  $i, j \in L$ , there exists a  $k \in \{0, \dots, |L|\}$  such that  $\Sigma_{ij}^k > 0$ . Note that  $\mathcal{P}(\Sigma) = \mathcal{P}(M(t))$  for all  $t \in \mathbb{N}$  since  $\sigma_{ij} > 0$  implies  $m_{ij}(t) \geq \delta$  for all  $t \in \mathbb{N}$  as shown above and, since every communication class of  $\Sigma$  is essential,  $m_{ij}(t) = 0$  if  $j \notin P_\Sigma(i)$ . Hence, for all  $L \in \mathcal{P}(\Sigma)$  and for all  $i, j \in L$  there exists a  $k \in \{0, \dots, |L|\}$  such that  $M(t+k, t)_{ij} > 0$  for all  $t \in \mathbb{N}$ .

Now, consider a sequence of time steps  $(t_s)_{s \in \mathbb{N}}$  such that  $t_0 = 0$  and  $t_{s+1} = t_s + \bar{L}$ , where  $\bar{L} := \max\{|L| : L \in \mathcal{P}(\Sigma)\}$ , and consider the sequence of accumulations  $(M(t_{s+1}, t_s))_{s \in \mathbb{N}}$ . By the rules of matrix multiplication, we get that for any two row stochastic  $A, B$  with a positive diagonal,  $(AB)_{ij} > 0$  if and only if  $A_{ij} > 0$  or  $B_{ij} > 0$ . Hence, for any  $L \in \mathcal{P}(\Sigma)$  and for all  $i, j \in L$ ,  $M(t+|L|, t)_{ij} > 0$  for all  $t \in \mathbb{N}$  since  $M(t)$  is row stochastic with a positive diagonal. Moreover,  $M(t+|L|, t)_{ij} = 0$  if  $j \notin P_\Sigma(i)$  since  $\mathcal{P}(\Sigma) = \mathcal{P}(M(t))$  for all  $t \in \mathbb{N}$ . Thus, for the accumulations  $M(t_{s+1}, t_s)$  it holds that  $M(t_{s+1}, t_s)_{ij} > 0$  if and only if  $j \in P_\Sigma(i)$ . In particular,  $M(t_{s+1}, t_s)$  is type-symmetric for all  $s \in \mathbb{N}$ .

For a non-negative matrix  $A$  let  $\min^+(A)$  denote the lowest positive entry of  $A$ . We have shown above that there exists a  $\delta > 0$  such that  $\sigma_{ij} > 0$  implies  $m_{ij}(t) \geq \delta$  for all  $t \in \mathbb{N}$ . Note that for any  $i, j \in L \in \mathcal{P}(\Sigma)$ , there exists a  $k \leq |L|$  and a sequence of dynasties  $(i_l)_{0 \leq l \leq k}$  with  $i_0 = i$  and  $i_k = j$  such that  $\sigma_{i_l, i_{l+1}} > 0$ , implying  $M(t+k, t)_{ij} \geq \prod_{l=0}^{k-1} m_{i_l, i_{l+1}}(t+l) \geq \delta^k$ . Thus, for the accumulations  $M(t_{s+1}, t_s)$  it holds that  $M(t_{s+1}, t_s)_{ij} \geq \delta^{t_{s+1}-t_s}$  if  $j \in P_\Sigma(i)$  and  $M(t_{s+1}, t_s)_{ij} = 0$  else. Hence,  $\min^+(M(t_{s+1}, t_s)) \geq \delta^{t_{s+1}-t_s} =: \delta^{|\bar{L}|}$ .

In summary, we have shown that the backward accumulation matrices  $(M(t_{s+1}, t_s))_{s \in \mathbb{N}}$  have a uniform lower bound of the positive entries  $\min^+(M(t_{s+1}, t_s)) \geq \delta^{|\bar{L}|}$ , are type-symmetric and have a strictly positive diagonal. By Lorenz (2005), Theorem 2, we get the desired result for the sequence  $(M(t_{s+1}, t_s))_{s \in \mathbb{N}}$ . Since  $\lim_{k \rightarrow \infty} \prod_{s=0}^k M(t_{s+1}, t_s) = \lim_{t \rightarrow \infty} M(t)$ , we also establish the statement of the Proposition.  $\square$

## References

Aleksynska, M. (2011). Civic participation of immigrants in Europe: Assimilation, origin, and destination country effects. *European Journal of Political Economy*, 27(3):566 – 585.

- Alesina, A. F., Giuliano, P., and Nunn, N. (2011). On the origins of gender roles: Women and the plough. *American Economic Review: Papers and Proceedings*, 101(3):499–503.
- Algan, Y. and Cahuc, P. (2010). Inherited trust and growth. *The American Economic Review*, 100:2060–2092.
- Bandura, A. (1977). *Social Learning Theory*. General Learning Press, New York, NY.
- Bisin, A. and Topa, G. (2003). Empirical models of cultural transmission. *J. Europ. Econ. Ass.*, 1(2-3):363–375.
- Bisin, A., Topa, G., and Verdier, T. (2009). Cultural transmission, socialization and the population dynamics of multiple-trait distributions. *Int. J. Econ. Theory*, 5(1):139–154.
- Bisin, A. and Verdier, T. (1998). On the cultural transmission of preferences for social status. *J. Pub. Econ.*, 70(1):75–97.
- Bisin, A. and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *J. Econ. Theory*, 97(2):298–319.
- Bisin, A. and Verdier, T. (2010). The economics of cultural transmission and socialization. In Benhabib, J.; Bisin, A. and Jackson, M., editors, *Handbook of Social Economics*, volume 1. North Holland.
- Botticini, M. and Eckstein, Z. (2005). Jewish occupational selection: Education, restrictions, or minorities? *The Journal of Economic History*, 65(04):922–948.
- Buechel, B., Hellmann, T., and Kloessner, S. (2014). Opinion dynamics and wisdom under conformity. Working Paper 51.2014, FEEM.
- Cavalli-Sforza, L. L. and Feldman, M. (1973). Cultural versus biological inheritance: phenotypic transmission from parents to children. *Am. J. Hum. Genet.*, 25:618–637.
- Cavalli-Sforza, L. L. and Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press, Princeton.
- Cygan-Rehm, K. (2012). Between here and there: Do immigrants follow their home countrys fertility norms? Technical report, University of Erlangen-Nuremberg.
- DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121.
- DeMarzo, P. M., Vayanos, D., and Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics*, 118(3):909–968.
- Doepke, M. and Zilibotti, F. (2008). Occupational choice and the spirit of capitalism. *The Quarterly Journal of Economics*, 123(2):747–793.
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2012). The intergenerational transmission of risk and trust attitudes. *Review of Economic Studies*, 79(2):645–677.

- Fernandez, R. and Fogli, A. (2006). Fertility: The role of culture and family experience. *Journal of the European Economic Association*, 4(2-3):552–561.
- Fernandez, R., Fogli, A., and Olivetti, C. (2004). Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics*, 119(4):1249–1299.
- Friedberg, S. H. and Insel, A. J. (1992). Convergence of matrix powers. *International Journal of Mathematical Education in Science and Technology*, 23(5):765–769.
- Fujimoto, T. and Ranade, R. R. (2004). Two characterizations of inverse–positive matrices: The Hawkins–Simon Condition and the Le Chatelier–Braun Principle. *Electronic Journal of Linear Algebra*, 11:59–65.
- Giavazzi, F., Petkov, I., and Schiantarelli, F. (2014). Culture: Persistence and Evolution. NBER Working Paper 20174.
- Golub, B. and Jackson, M. O. (2010). Naïve learning in social networks and the wisdom of crowds. *Amer. Econ. J.: Microecon.*, 2(1):112–49.
- Grusec, J. E. (2002). *Handbook of Parenting: Practical Issues in Parenting*, chapter Parental Socialization and Children’s Acquisition of Values, pages 153–167. Lawrence Erlbaum Associates.
- Guiso, L., Sapienza, P., and Zingales, L. (2008). Alfred Marshall lecture social capital as good culture. *Journal of the European Economic Association*, 6(2-3):295–320.
- Horn, R. A. and Johnson, C. R. (2010). *Matrix analysis*. Cambridge Univ. Press, Cambridge.
- Jackson, M. O. (2008). *Social and Economic Networks*. Princeton University Press.
- Kalai, A. and Kalai, E. (2001). Strategic polarization. *J. Math. Psychol.*, 45:656–663.
- Lorenz, J. (2005). A stabilization theorem for dynamics of continuous opinions. *Physica A: Statistical Mechanics and its Applications*, 355(1):217–223.
- Lorenz, J. (2006). Convergence of products of stochastic matrices with positive diagonals and the opinion dynamics background. In Commault, C. and Marchand, N., editors, *Positive Systems*, volume 341 of *Lecture Notes in Control and Information Sciences*, pages 209–216. Springer Berlin / Heidelberg.
- Martinez, S., Michon, G., and Martin, J. S. (1994). Inverse of strictly ultrametric matrices are of Stieltjes type. *SIAM J. Matrix Anal. Appl.*, 15(1):98–106.
- Nabben, R. and Varga, R. S. (1993). On symmetric ultrametric matrices. In Reichel, L., Ruttan, A., and Varga, R. S., editors, *Numerical Linear Algebra*. Walter de Gruyter, New York.
- Nabben, R. and Varga, R. S. (1994). A linear algebra proof that the inverse of a strictly ultrametric matrix is a strictly diagonally dominant Stieltjes matrix. *SIAM J. Matrix Anal. Appl.*, 15(1):107–113.

- Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in africa. *American Economic Review*, 101(7):3221–3252.
- Ostrowski, A. M. (1959). A quantitative formulation of Sylvester’s law of inertia. *Proc Natl Acad Sci U S A*, 45(5):740–744.
- 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. (2014). Socialization networks and the transmission of interethnic attitudes. *Journal of Economic Theory*, 150:583–610.
- Pichler, M. M. (2010). The economics of cultural formation of preferences. Working Papers 431, Bielefeld University, Institute of Mathematical Economics.
- Tabellini, G. (2010). Culture and institutions: economic development in the regions of europe. *Journal of the European Economic Association*, 8(4):677–716.
- Vaughan, D. (2010). Conformity and the Transmission of Continuous Cultural Traits. Departmental Working Papers 2011-01, Department of Economics, New York University.
- Voigtländer, N. and Voth, H.-J. (2012). Persecution perpetuated: The Medieval origins of anti-Semitic violence in Nazi Germany. *The Quarterly Journal of Economics*, 127(3):1339–1392.