Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 198 (2021) 105373

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Cultural transmission with incomplete information ☆

Sebastiano Della Lena [a,*], Fabrizio Panebianco [b]

[a] *Department of Economics, University of Antwerp, Prinsstraat 13 2000 Antwerpen, Belgium*
[b] *Department of Economics and Finance, Università Cattolica del Sacro Cuore, Milano, Italy*

Received 7 June 2021; final version received 13 October 2021; accepted 18 October 2021
Available online 21 October 2021

Abstract

This paper introduces incomplete information into the standard cultural transmission framework (Bisin and Verdier, 2001). We consider parents having incomplete information about population shares and about the efficiency of their transmission technology. We show that conjectures about population shares are the key determinants of long-run population configurations. Namely, if these conjectures are positively or mildly negatively biased, there is always long-run cultural heterogeneity. If, instead, they are strongly negatively biased, long-run cultural homogeneity is displayed. We also find that, depending on the properties of conjectures about efficiency of parental transmission technology, standard cultural substitution may not hold. Notably, differently from the literature, cultural substitution, when displayed, does not guarantee long-run cultural heterogeneity. Then, considering parents who, before socializing children, experiment to acquire information, we show that they may not be able to disentangle the impact of the efficiency of their transmission technology from that of population share. Thus, parents generally fail to learn about the unknowns. We conclude the paper with a brief discussion about how cultural leaders may negatively bias conjectures about population shares and foster cultural homogeneity.

JEL classification: D10; D80; J10; Z10

☆ We are grateful to Pierpaolo Battigalli, Alberto Bisin, and Alessandro Pavan for helpful comments and suggestions. We also thank Sergio Currarini, Simon Weidenholzer, Thierry Verdier, and anonymous referees. Sebastiano Della Lena gratefully acknowledges funding from FWO-foundation for the project *"Diffusion of Misinformation in Social Networks"* contract n. 1258321N. Fabrizio Panebianco gratefully acknowledges funding from Spanish Ministry of Economia y Competitividad project ECO2017-87245-R and the ERC grant 324219.

* Corresponding author.
E-mail addresses: dellalena.sebastiano@uantwerpen.be (S. Della Lena), fabrizio.panebianco@unicatt.it (F. Panebianco).

https://doi.org/10.1016/j.jet.2021.105373

Keywords: Cultural transmission; Incomplete information; Group size misperception; Parental efficacy; Cultural leaders

1. Introduction

This paper introduces *incomplete information* into the standard probabilistic cultural transmission framework (Bisin and Verdier, 2001). We consider incomplete information along two dimensions: the share of different cultural groups within society; and the level of efficiency of parental transmission technology. We show how incomplete information affects both parental socialization efforts and long-run population dynamics, and that it can revert standard results of the persistence of cultural heterogeneity in the society. We also show that parents, even if they acquire information during the parenting process, generally fail to learn the unknowns.

In the literature (e.g., Bisin and Verdier, 2001, 2011), the transmission of cultural traits from one generation to the next is the result of *vertical* (or direct) *socialization* — i.e., purposeful socialization by the family — and *oblique socialization* — i.e., socialization by the society.[1] In this paper, we introduce incomplete information as affecting both vertical and oblique socialization.

Consider first vertical socialization. Parents exert some effort to socialize their offspring to their own cultural traits. The choice of effort and the success of vertical socialization depend on the efficiency of the effort exerted. This efficiency is known in social psychology as **parental efficacy** (Belsky et al., 1984; Van Bakel and Riksen-Walraven, 2002; Belsky and Jaffee, 2006). In economic models of cultural transmission, this dimension is often disregarded and parents are generally assumed to be fully efficient and aware that they are so.[2] However, we argue that parental efficacy is hardly objectively measurable and, thus, parents merely conjecture theirs through self-assessments based on their available information. For these reasons, we refer to the theory of *self-efficacy* (Bandura, 1982, 1993; Bandura et al., 1999) and we talk about **conjectured parental efficacy**, that is, parents' conjectures about their own parental efficacy.[3]

The second dimension into which we introduce incomplete information is oblique socialization. When determining how much effort to exert on socialization, parents also take into account the composition of the society affecting offspring. As standard, we assume that the society is partitioned into two cultural groups and, differently from the literature, parents do not have complete information about the population shares. There is indeed evidence that conjectures about

[1] The study of cultural transmission stems from Cavalli-Sforza and Feldman (1981). In economic literature, starting from Bisin and Verdier (2001), the study of cultural transmission developed in several directions. Among others, Büchel et al. (2014) and Panebianco (2014) consider the role of network structure in the transmission of continuous traits, Vaughan (2013) introduces the role of peers in horizontal socialization, Bisin and Verdier (2017) analyze the joint evolution of culture and institutions, Cheung and Wu (2018) generalize the standard framework with probabilistic transmission to continuous traits, Spiro (2020) considers the multigenerational transmission of culture, and Bisin et al. (2004), Tabellini (2008), and Della Lena and Dindo (2019) study the role of strategic environments in the transmission of traits.

[2] Note that there is a growing literature studying the role of parenting styles and their impact on children's outcomes, for example, Doepke and Zilibotti (2017, 2019).

[3] This concept in social psychology is mostly known as *perceived parental efficacy*. de Montigny and Lacharité (2005) define perceived parental efficacy as "beliefs or judgments parents hold of their capabilities to organize and execute a set of tasks related to parenting a child". In the literature, perceived parental efficacy is also known as *perceived parental self-efficacy* or, simply, *parental self-efficacy*.

population shares can be biased by (fake) news, by the *salience* of cultural traits in the public debate (Alesina et al., 2018), or also by the declarations of cultural and political leaders.[4]

In Section 2, we show the impact of possibly wrong conjectures about the unknowns both on parental choices and long-run cultural dynamics. At first, we characterize subjectively optimal socialization efforts, showing that the more parents believe themselves effective in parenting, the higher incentives they have to exert effort, whereas the more parents believe their own trait is present in society, the less they are incentivized to socialize children by themselves. We show that, under incomplete information, the probability of a child acquiring the parental trait can be higher than it is under complete information, and we provide necessary and sufficient conditions for this to happen. Nevertheless, the subjectively optimal effort is objectively suboptimal and the associated utility loss is proportional to the difference between the effort exerted and the optimal effort under complete information.

In Section 2.1, we study the implications of incomplete information for population dynamics. We show that, depending on how much and in which direction the conjectures about population shares are biased, different long-run cultural configurations may arise. In particular, if the biases are either positive or mildly negative, stable cultural heterogeneity is observed in the long run, as under complete information (Bisin and Verdier, 2001). Instead, if conjectures about population shares are strongly negatively biased, the socialization efforts of an extreme majority are strong enough to dominate minority ones and, thus, extreme steady states become stable and long-run cultural homogeneity is observed.

In Section 2.2, we analyze how socialization efforts react to variation in *actual* population shares and we provide conditions for *cultural substitution* or *complementarity*.[5] Notably, a standard result when traits are probabilistically transmitted from one generation to the next (Bisin and Verdier, 2001; Cheung and Wu, 2018) is that a sufficient condition for long-run cultural heterogeneity is socialization efforts displaying *cultural substitution*. Conversely, in our model, we show that under incomplete information cultural substitution does not necessarily lead to long-run cultural heterogeneity, and that cultural substitution (complementarity) can coexist with long-run cultural homogeneity (heterogeneity).

In Section 3, we consider how parents may exploit parenting time to acquire information about the unknowns. Indeed, given the nature of cultural transmission, parents constantly interact and share information with their own children. For this reason, we assume that parenting time is constituted by a period of **active learning** — in which parents acquire information through experimentation — and an *actual* **socialization period** — in which the traits are acquired. During active learning, each parent exerts one or more experimentation efforts and, after each, receives messages from the child about the probability of acquiring the parental trait. Based on the messages received after each experimentation effort, the parent updates their own conjectures about the unknowns. During the socialization period, parents continuously exert the optimal socialization effort given the subjective conjectures formed during active learning, and children acquire a trait with a probability depending only on this effort.

In Section 3.1, we consider the case in which the parent uses the conjectures updated following a *single* experimentation effort to actually socialize the child. Conversely, motivated by the fact

4 As a matter of example, in 2016 *The Economist* analyzed the perceptions of European citizens have about the share of the Muslim population in their own countries, showing that these perceptions are extremely biased (*The Economist*, “Islam in Europe: Perception and reality”).

5 A socialization effort displays cultural substitution (complementarity) if it negatively (positively) reacts to an increase in the share of own trait in society.

that the posterior conjectures after a single experimentation effort are generally different from the prior conjectures, in Section 3.2 we consider the possibility that parents go on experimenting and updating conjectures until the process converges.[6] In both cases, we first provide results about the locus of conjectures compatible with agents' active learning, and then we characterize the optimal socialization efforts. We find that parents generally fail to learn the two unknowns and, thus, they exert suboptimal socialization efforts. This is due to the fact that they are not able to disentangle the impact of their parental efficacy from that of population shares on the observed outcome of their experimentation. Notably, increasing the number of messages parents receive and the number of experimentation efforts does not guarantee that parents exert a socialization effort closer to the objectively optimal one. We also discuss, with the help of two examples, how conjectures endogenously formed through active learning relate to cultural complementarity and substitution results discussed in Section 2.2.

In Section 4, we provide a discussion of the case in which conjectures about population shares may be biased by cultural leaders.[7] We consider both utilitarian and identitarian leaders —i.e., leaders maximizing the utility of agents and the share of own group population in the next period, respectively. We show that utilitarian leaders always induce correct conjectures, whereas identitarian ones always induce negatively biased conjectures. Coherently with our results about population dynamics in Section 2.1, the presence of identitarian leaders pushes towards cultural homogeneity by means of a tightening of agents' socialization efforts. This may help explain how cultural conflicts between groups may lead to instability of the heterogeneous societies and, thus, the extinction of weak traits in the long run.[8]

In Section 5, we conclude the paper and in the Appendix we provide the proofs of the results.

2. The model

Consider a society composed by a continuum of agents and partitioned into two cultural groups. Each agent i is endowed with a cultural trait from the set $\mathcal{C} := \{A, B\}$. With a little abuse of notation, we denote by I both the generic trait $I \in \mathcal{C}$ and the set of all agents displaying that trait, i.e., cultural group I. At each time $t \in \mathbb{N}_0$ the fraction of agents with the trait $I \in \mathcal{C}$ is $q_t^{\mathrm{I}} \in [0, 1]$ and each agent reproduces asexually, giving birth to one child. Traits pass from one generation to the next according to a cultural transmission process at the end of which children acquire traits, become adults, and the process starts again.

During the cultural transmission process, each parent $i \in I$ directly socializes their own child to trait I with an endogenous probability $d_t^i \in [0, 1]$, which depends on a costly **socialization effort** $\tau_t^i \in [0, 1]$, and on an idiosyncratic technology, $\alpha^i \in \mathbb{R}_+$, capturing the **parental efficacy** with which each parent transmits own trait. In each cohort, the parental efficacies of agents in I are $i.i.d.$ according to a time-independent measure with mean $\alpha^{\mathrm{I}} := \mathbb{E}[\alpha^i]$. We assume, as a simple generalization of Bisin and Verdier (2001), that $d_t^i := d(\alpha^i, \tau_t^i) = \min\{1, \alpha^i \tau_t^i\}$.[9]

[6] We discuss in the paper that each steady state of such updating process is equivalent to a *selfconfirming equilibrium* (Battigalli, 1987; Fudenberg and Levine, 1993; Battigalli et al., 2015).

[7] For previous works about cultural leaders, refer to Nteta and Wallsten (2012); Acemoglu and Jackson (2014); Verdier and Zenou (2015); Prummer and Siedlarek (2017); Verdier and Zenou (2018).

[8] This result is complementary to that Prummer and Siedlarek (2017), where the authors discuss cultural leaders as the cause of different traits coexisting in a framework with continuous traits, deterministic transmission, and complete information.

[9] Bisin and Verdier (2001) in their leading example implicitly assume $\alpha^i = 1$, that is $d_t^i = \tau_t^i$, whereas, in another example, "It takes a village to raise a child", they consider the case of $\alpha^i = q_t^{\mathrm{I}}$.

This process is known as **vertical socialization**. If vertical socialization fails, which occurs with probability $1 - d_t^i$, the child is indirectly socialized by picking the trait of a role model randomly chosen from the population. This process is known as **oblique socialization**. Therefore, for each $I \in \mathcal{C}$ and each parent $i \in I$, the probability that their own child is socialized to their own trait is

$$p_t^{i\mathrm{I}} := p(\tau_t^i; \alpha^i, q_t^{\mathrm{I}}) = \underbrace{\alpha^i \tau_t^i}_{vertical\ soc.} + \underbrace{(1 - \alpha^i \tau_t^i) q_t^{\mathrm{I}}}_{oblique\ soc.}. \tag{1}$$

Given $J \neq I$, we also define $p_t^{i\mathrm{J}} := 1 - p_t^{i\mathrm{I}}$, that is, the probability that i's own child is socialized to the other trait.

Incomplete information We consider that each parent $i \in I$, of each group $I \in \mathcal{C}$, has incomplete information about their own parental efficacy α^i and about population share q_t^{I}. Let $\mu_t^i := (\mu_{\alpha^i,t}^i, \mu_{q_t^{\mathrm{I}},t}^i) \in \Delta(\mathbb{R}_+) \times [0,1]$ be the vector of i's conjectures about unknowns.[10] In detail, each i has a probabilistic conjecture about own parental efficacy, $\mu_{\alpha^i,t}^i \in \Delta(\mathbb{R}_+)$, with subjective expectation $\hat{\alpha}_t^i := \mathbb{E}_{\mu_t^i}^i[\alpha^i]$, which we refer to as **conjectured parental efficacy**. Conversely, the conjecture about own group population share is assumed to be deterministic, namely, $\mu_{q_t^{\mathrm{I}},t}^i = \delta_{\hat{q}_t^i}$ with $\mathbb{E}_{\delta_{\hat{q}_t^i}}^i[q_t^{\mathrm{I}}] = \hat{q}_t^i \in [0,1]$.[11] We further assume that $\hat{q}_t^i$ may depend on population share q_t^{I} and, in such a case, to be non-decreasing, namely $\frac{\partial \hat{q}_t^i}{\partial q_t^{\mathrm{I}}} \geq 0$.

Conjectures about parental efficacy and population shares, μ_t^i, induce, through (1), a probabilistic conjecture about transition probability, $\pi_t^i \in \Delta([0,1])$. We define as **conjectured transition probability** the expectation about the transition probability,

$$\hat{p}_t^{i\mathrm{I}} := \mathbb{E}_{\pi_t^i}^i[p_t^{i\mathrm{I}}] = \hat{\alpha}_t^i \tau_t^i + (1 - \hat{\alpha}_t^i \tau_t^i) \hat{q}_t^i. \tag{2}$$

Let also define $\hat{p}_t^{i\mathrm{J}} := 1 - \hat{p}_t^{i\mathrm{I}}$.

As standard, we assume that parents prefer that a child has their own, rather than a different trait. In detail, for each $i \in I$, let $V^{i\mathrm{I}} \in [0,1]$ and $V^{i\mathrm{J}} \in [0,1]$ be the exogenous utilities the parent i gets from having a child with trait I and J, respectively. We model parents' preferences to be group dependent. Then, for each $i \in I$, $V^{i\mathrm{I}} = V^{\mathrm{II}}$ and $V^{i\mathrm{J}} = V^{\mathrm{IJ}}$, where $V^{\mathrm{II}} > V^{\mathrm{IJ}}$. Assuming quadratic socialization costs, each parent $i \in I$, chooses an effort $\tau_t^i \in [0,1]$ to maximize their own subjective expected utility

$$\mathbb{E}_{\pi_t^i}^i[u(\tau_t^i)] = \hat{p}_t^{i\mathrm{I}} V^{\mathrm{II}} + \hat{p}_t^{i\mathrm{J}} V^{\mathrm{IJ}} - \frac{1}{2}(\tau_t^i)^2. \tag{3}$$

Then, for each $i \in I$, the subjectively optimal socialization effort is given by

$$\tau_t^i = \hat{\alpha}_t^i (1 - \hat{q}_t^i) \Delta V^{\mathrm{I}}, \tag{4}$$

where $\Delta V^{\mathrm{I}} := V^{\mathrm{II}} - V^{\mathrm{IJ}}$ represents the *cultural intolerance* of a parent with trait I, namely, the relative value of a child with the same trait as the parent compared with a child with a

[10] We refer to Section 3 for an analysis of the process that leads agents to form conjectures about unknowns.

[11] We choose to keep conjectures about population share deterministic because, as discussed in the introduction, conjectured parental efficacy usually stems from a self-assessment and, thus, agents are more susceptible of being uncertain about $\hat{\alpha}_t^i$ rather than $\hat{q}_t^i$, which is more likely to be affected by external sources of information. We can provide results also for probabilistic conjectures about population shares q_t^{I}. In that case, we should take into account the covariance between the conjectures; however, results do not qualitatively change.

different trait. From equation (4) we see that the optimal socialization effort is increasing in the conjectured parental efficacy,[12] and decreasing in the conjecture about the share of the parents' cultural group within the society.

In the benchmark case of complete information, the socialization effort is computed using correct conjectures and we define it as $\tau^i_{bv,t} := \alpha^i(1-q^{\mathrm{I}}_t)\Delta V^{\mathrm{I}}$, where the subscript stands for Bisin and Verdier (2001). Then, $\tau^i_{bv,t}$ induces a transition probability $p^{i\mathrm{I}}_{bv,t}$. We can now compare, from equation (1), the transition probabilities under complete and incomplete information. For each $I \in \mathcal{C}$, we denote by $b^i_t := \hat{q}^i_t - q^{\mathrm{I}}_t$ the idiosyncratic bias of each agent $i \in I$ in conjectures about own population share.

Proposition 1. *For each $i \in I$ and $I \in \mathcal{C}$, $p^{i\mathrm{I}}_t > p^{i\mathrm{I}}_{bv,t}$ if and only if $b^i_t < (1-q^{\mathrm{I}}_t)\left(\frac{\hat{\alpha}^i_t - \alpha^i_t}{\hat{\alpha}^i_t}\right)$.*

The main message of this proposition is that if the bias in conjectures about population shares is low, then τ^i_t is high and, consequently, the transition probability is more likely to be higher than under complete information. For the intuition of this result, note first that, by equation (4), the effects on τ^i_t of $\hat{\alpha}^i_t$ and $\hat{q}^i_t$ are positive and negative, respectively. Thus, if $b^i_t < 0$ and $\hat{\alpha}^i_t > \alpha^i$, it trivially follows that parent i exerts a lower effort than under complete information and, consequently, the transition probability is lower. On the contrary, if $b^i_t > 0$ and $\hat{\alpha}^i_t < \alpha^i$, the socialization effort and the transition probability are lower than under complete information. Lastly, if the two inequalities have the same sign, the effects go in opposite directions and the ordering of $p^{i\mathrm{I}}_t$ and $p^{i\mathrm{I}}_{bv,t}$ depends on the relative strength of the two effects.

Objective expected utility loss Given incomplete information about parental efficacy and population shares, parents may not be able to exert the objectively optimal socialization effort and, in that case, they may gain a lower utility than under complete information. For each $I \in \mathcal{C}$ and $i \in I$, let $\mathbb{E}_{p^{i\mathrm{I}}}[u(\tau^i_t)]$ be i's objective expected utility given τ^i_t, and let $\mathbb{E}_{p^{i\mathrm{I}}_{bv,t}}[u(\tau^i_{bv,t})]$ be the i's objective expected utility given $\tau^i_{bv,t}$.

Proposition 2. *For each $I \in \mathcal{C}$ and $i \in I$, the objective expected individual loss is given by*

$$\Delta U^i_t := \mathbb{E}_{p^{i\mathrm{I}}_t}[u(\tau^i_t)] - \mathbb{E}_{p^{i\mathrm{I}}_{bv,t}}[u(\tau^i_{bv,t})] = \frac{1}{2}(\tau^i_t - \tau^i_{bv,t})^2. \tag{5}$$

We can see that the individual loss induced by incomplete information quadratically depends on the distance between subjective and objective optimal socialization efforts. Note that parents, who suffer from incomplete information, cannot compute the objective expected utilities $\mathbb{E}_{p^{i\mathrm{I}}_t}[u(\tau^i_t)]$ and $\mathbb{E}_{p^{i\mathrm{I}}_{bv,t}}[u(\tau^i_{bv,t})]$. These utilities can be computed by a fully informed policy maker (e.g., see the discussion about utilitarian cultural leaders in Section 4).

12 This result is in line with the social psychology literature, in which, as outlined in Coleman and Karraker (2000), a high level of conjectured parental efficacy has been found to predict a high level of parental effort and performance, for example, responsiveness to children's needs (Donovan and Leavitt, 1985; Donovan et al., 1997; Unger and Wandersman, 1985), engagement in direct parenting interactions (Mash and Johnston, 1983), and active parental coping strategies (Wells-Parker et al., 1990).

2.1. Population dynamics

We now consider how incomplete information impacts the population dynamics. In particular, we show that we reach either long-run cultural homogeneity or heterogeneity, depending on the properties of agents' conjectures.

The population dynamics averages out the effects of the efforts of all agents. Thus, we first define and compute the average probability of vertical socialization of agents belonging to group I, given optimal efforts in equation (4), as

$$\begin{aligned} d_t^{\mathrm{I}} := \mathbb{E}[d_t^i] &= \alpha^{\mathrm{I}}\tau_t^{\mathrm{I}} + cov[\alpha^i, \tau_t^i] \\ &= \alpha^{\mathrm{I}}\Big(\hat{\alpha}_t^{\mathrm{I}}(1-\hat{q}_t^{\mathrm{I}})\Delta V^{\mathrm{I}}\Big) + (1-\hat{q}_t^{\mathrm{I}})\Delta V^{\mathrm{I}} cov[\alpha^i, \hat{\alpha}_t^i] \\ &= \omega_t^{\mathrm{I}}(1-\hat{q}_t^{\mathrm{I}}), \end{aligned} \tag{6}$$

where $\hat{\alpha}_t^{\mathrm{I}} := \mathbb{E}[\hat{\alpha}_t^i]$ and $\omega_t^{\mathrm{I}} := \mathbb{E}[\alpha^i\hat{\alpha}_t^i]\Delta V^{\mathrm{I}}$. Recall that with complete information $d_t^{\mathrm{I}} = \alpha^{\mathrm{I}2}(1 - q_t^{\mathrm{I}})\Delta V^{\mathrm{I}}$. Conversely, with incomplete information the determinants of d_t^{I} are: (i) average conjectures about population shares, $\hat{q}_t^{\mathrm{I}}$, and (ii) average actual and conjectured parental efficacies, α^{I} and $\hat{\alpha}_t^{\mathrm{I}}$.

For each $I \in \mathcal{C}$, let us define the average transition probability as $p_t^{\mathrm{II}} := \mathbb{E}[p_t^{i\mathrm{I}}]$, so that the equation describing the dynamics of population share q_t^{I} is given by

$$\begin{aligned} q_{t+1}^{\mathrm{I}} &= q_t^{\mathrm{I}} p_t^{\mathrm{II}} + (1-q_t^{\mathrm{I}})(1-p_t^{\mathrm{JJ}}) \\ &= q_t^{\mathrm{I}}[1 + (d_t^{\mathrm{I}} - d_t^{\mathrm{J}})(1-q_t^{\mathrm{I}})]. \end{aligned}$$

Using, as standard, a continuous-time approximation we get:

$$\dot{q}_t^{\mathrm{I}} = q_t^{\mathrm{I}}(1-q_t^{\mathrm{I}})(d_t^{\mathrm{I}} - d_t^{\mathrm{J}}). \tag{7}$$

Note that $q^{\mathrm{I}} = 0$ and $q^{\mathrm{I}} = 1$ are always steady states of (7), whereas interior steady states, if they exist, satisfy $(d_t^{\mathrm{I}} - d_t^{\mathrm{J}}) = 0$.

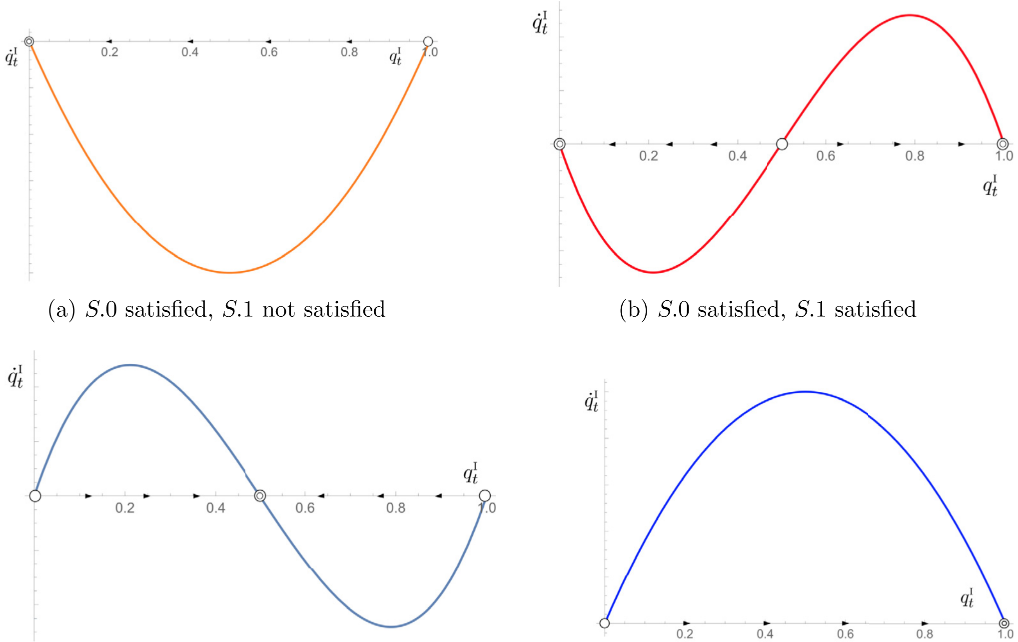
The next proposition discusses the stability of the extreme points, the characterization of the interior (i.e., polymorphic) equilibria, and their existence, uniqueness, and stability properties. Analyzing the population dynamics, we find it convenient to make explicit the dependence of the bias from the population shares. Fixing $I \in \mathcal{C}$, for each $i \in I$, we write the bias as $b_t^i(q_t^{\mathrm{I}}) = \hat{q}_t^i - q_t^{\mathrm{I}}$ and, thus, its average in population I as $b_t^{\mathrm{I}}(q_t^{\mathrm{I}}) = \mathbb{E}[b_t^i(q_t^{\mathrm{I}})]$. Similarly, for $j \in J \neq I$, let $b_t^j(q_t^{\mathrm{I}}) = \hat{q}_t^j - (1-q_t^{\mathrm{I}})$ and $b_t^{\mathrm{J}}(q_t^{\mathrm{I}}) = \mathbb{E}[b_t^j(q_t^{\mathrm{I}})]$. Notice that, at $q_t^{\mathrm{I}} = 0$ the bias of each $i \in I$ must be non-negative (i.e., $b_t^i(0) \geq 0$), and the bias of each agent $j \in J$ must be non-positive (i.e., $b_t^j(0) \leq 0$); for the same reason, the opposite happens at $q_t^{\mathrm{I}} = 1$. Therefore, looking at the average biases, it always holds that $b_t^{\mathrm{I}}(0) \geq 0$, $b_t^{\mathrm{J}}(0) \leq 0$, $b_t^{\mathrm{I}}(1) \leq 0$, and $b_t^{\mathrm{J}}(1) \geq 0$.

Proposition 3. *Consider the dynamics in (7). Then, $q^{\mathrm{I}} = 0$ and $q^{\mathrm{I}} = 1$ are always steady states and, given average steady state conjectures $(\hat{\alpha}^{\mathrm{I}}, \hat{q}^{\mathrm{I}})_{I \in \mathcal{C}}$:*

S.0 *: $q^{\mathrm{I}} = 0$ is stable if and only if $b^{\mathrm{I}}(0) > 1 - |b^{\mathrm{J}}(0)|\frac{\omega^{\mathrm{J}}}{\omega^{\mathrm{I}}}$;*
S.1 *: $q^{\mathrm{I}} = 1$ is stable if and only if $|b^{\mathrm{I}}(1)| > \big(1 - b^{\mathrm{J}}(1)\big)\frac{\omega^{\mathrm{J}}}{\omega^{\mathrm{I}}}$.*

Each interior equilibrium of the dynamics is a $q^{\mathrm{I}*} \in (0, 1)$ such that

$$q^{\mathrm{I}*} = \frac{\omega^{\mathrm{I}}}{\omega^{\mathrm{I}} + \omega^{\mathrm{J}}}\big(1 - b^{\mathrm{I}}(q^{\mathrm{I}*})\big) + \frac{\omega^{\mathrm{J}}}{\omega^{\mathrm{I}} + \omega^{\mathrm{J}}} b^{\mathrm{J}}(q^{\mathrm{I}*}). \tag{8}$$



(a) $S.0$ satisfied, $S.1$ not satisfied
(b) $S.0$ satisfied, $S.1$ satisfied
(c) $S.0$ not satisfied, $S.1$ not satisfied
(d) $S.0$ not satisfied, $S.1$ satisfied

Fig. 1. Population dynamics $\dot{q}_t^{\mathrm{I}}$ as dependent on conditions $S.0$ and $S.1$ in Proposition 3, when there exists at most one of interiors steady state $q^{\mathrm{I}*}$.

If $q^{\mathrm{I}} = 0$ and $q^{\mathrm{I}} = 1$ are either both stable or both unstable, and if $\hat{q}_t^{\mathrm{I}}$ and $\hat{q}_t^{\mathrm{J}}$ are one concave and one convex in q_t^{I}, then $q^{\mathrm{I}*} \in (0, 1)$ is unique.

Proposition 3 provides the conditions on the population shares conjectures for the stability of extreme steady states, $q^{\mathrm{I}} = 0$ and $q^{\mathrm{I}} = 1$ (conditions $S.0$ and $S.1$, respectively), and also for having a unique (stable) interior steady state, $q^{\mathrm{I}*} \in (0, 1)$. Note that, having both $q^{\mathrm{I}} = 0$ and $q^{\mathrm{I}} = 1$ unstable (stable) guarantees the stability (instability) of $q^{\mathrm{I}*}$, when unique. For $q^{\mathrm{I}*}$ to be unique it is enough that the concavity of (average) conjectures in *own* population share is the same for both groups and it does not change with q^{I}.[13] Moreover, in equation (8) we provide a characterization of any possible $q^{\mathrm{I}*} \in (0, 1)$.

Fig. 1 shows the possible classes of population dynamics when the conditions for the uniqueness of $q^{\mathrm{I}*}$ are satisfied. We refer to Fig. 1 in the following discussion of the stability of extreme points $q = 0$ and $q = 1$ and long-run configurations.

For the intuition of the results, consider the stability properties of the extreme steady state $q^{\mathrm{I}} = 0$. In the neighborhood of this point, I agents can only have non-negative biases (i.e., $b^{\mathrm{I}}(0) \geq 0$ always). For this reason, agents in I choose an average effort that is lower than (or equal to) the average objectively optimal one, τ_{bv}^{I}. This reduces the probability that, in the next period, children would be of I type and, consequently, it negatively affects $\dot{q}_t^{\mathrm{I}}$. In a similar way, J agents

[13] Throughout the paper, we use as examples both exogenous conjectures about population shares (Example 1) and conjectures linear in own population shares (Example 2 and Example 3). Both cases satisfy the conditions for unique interior steady state.

cannot overestimate their population share, $b^{\mathrm{J}}(0) \leq 0$, producing an average higher (or equal) effort than the objectively optimal one and this again negatively affects $\dot{q}_t^{\mathrm{I}}$. Thus, as stated in $S.0$, if the minority has an average bias that is too large relatively to the majority (i.e., $b^{\mathrm{I}}(0) > 1 - |b^{\mathrm{J}}(0)|\frac{\omega^{\mathrm{J}}}{\omega^{\mathrm{I}}}$), then I trait does not survive for small q_t^{I} and $q^{\mathrm{I}} = 0$ is a stable steady state (Figs. 1a and 1b). A similar argument explains the stability of $q^{\mathrm{I}} = 1$ (Figs. 1b and 1d). Lastly, if the average biases of agents belonging to I and J are both non-negative, then $S.0$ and $S.1$ are never satisfied, thus, as in the standard framework (Bisin and Verdier, 2001), $q^{\mathrm{I}} = 0$ and $q^{\mathrm{I}} = 1$ are both unstable and there exists an interior stable steady state (Fig. 1c). The same holds if the average biases are mildly negative, because the subjectively optimal socialization efforts are close to the objectively optimal ones and, thus, the dynamics is not qualitatively affected by incomplete information.

The main intuition of Proposition 3 is that incomplete information about population shares can have a severe impact on population dynamics, because it makes the extreme steady states more stable. Indeed, agents belonging to a very tiny minority, if they have wrong conjectures about population shares, are most likely to overestimate own presence in the society. This reduces their socialization efforts and, depending on the strength of the biases, can make extreme points ($q^{\mathrm{I}} = 0$ and $q^{\mathrm{I}} = 1$) stable. Therefore, compared to complete information, incomplete information about population shares opens the door to the extinction of the cultural minority and, thus, to long-run cultural homogeneity. Note that results qualitatively depend only on incomplete information about population shares and that all the long-run configurations previously discussed can exist *independently* of the assumptions about parental efficacy (e.g., if parental efficacies are homogeneous and common knowledge, $S.0$ and $S.1$ become $b^{\mathrm{I}}(0) > 1 - |b^{\mathrm{J}}(0)|$ and $b^{\mathrm{I}}(1) > 1 - b^{\mathrm{J}}(1)$, respectively).

Proposition 3 and the previous discussion highlight the role of incomplete information about population shares for the long-run dynamics. We now focus on the specific role of incomplete information about parental efficacy. To do that, we consider the particular case in which parents know the true value of q^{I} but they still have uncertainty about α^i.

Corollary 3.1. *Consider the dynamics in (7), and, for each $I \in \mathcal{C}$, let $\hat{q}_t^{\mathrm{I}} = q_t^{\mathrm{I}}$. Then, the unique interior steady state is $q^{\mathrm{I}*} = \frac{\omega^{\mathrm{I}}}{\omega^{\mathrm{I}} + \omega^{\mathrm{J}}} \in (0, 1)$, which is also stable.*

Corollary 3.1 is a particular case of Proposition 3 when conjectures about population shares are correct on average. Crucially, incomplete information about parental efficacy in itself does not qualitatively change results about long-run dynamics, if compared to complete information — which displays the same dynamics as in Fig. 1c. Indeed, in both cases, there exists a unique interior stable steady state. Note that, when there is complete information about parental efficacies and they are homogeneous among groups, $q^{\mathrm{I}*} = \frac{\omega^{\mathrm{I}}}{\omega^{\mathrm{I}} + \omega^{\mathrm{J}}} = \frac{\Delta V^{\mathrm{I}}}{\Delta V^{\mathrm{I}} + \Delta V^{\mathrm{J}}}$.

2.2. Comparative statics

We now analyze how incomplete information has an impact on the way in which equilibrium socialization efforts react to change in population shares. We say that socialization effort displays **cultural substitution** if $\frac{\partial \tau_t^i}{\partial q_t^{\mathrm{I}}} < 0$, and **cultural complementarity** if $\frac{\partial \tau_t^i}{\partial q_t^{\mathrm{I}}} > 0$.

Proposition 4. *Fix $i \in I$, and consider the optimal socialization effort in equation (4). Then,*

(i) *if* $\frac{\partial \hat{\alpha}_t^i}{\partial q_t^{\mathrm{I}}} \leq 0$, τ_t^i *displays cultural substitution;*

(ii) *if* $\frac{\partial \hat{\alpha}_t^i}{\partial q_t^{\mathrm{I}}} > 0$, τ_t^i *displays complementarity if and only if* $\frac{\partial \hat{q}_t^i}{\partial q_t^{\mathrm{I}}} < \frac{(1-\hat{q}_t^i)}{\hat{\alpha}_t^i} \frac{\partial \hat{\alpha}^i}{\partial q_t^{\mathrm{I}}}$.

To understand this result, recall that $\tau_t^i = \hat{\alpha}_t^i (1 - \hat{q}_t^i) \Delta V^{\mathrm{I}}$ and that we assumed $\hat{q}_t^i$ to be non-decreasing in q_t^{I}. Thus, an increase in q_t^{I} makes the effort decreasing through its effect on $\hat{q}_t^i$. Consequently, if the conjectured parental efficacy $\hat{\alpha}_t^i$ is independent of, or decreasing on, population shares, cultural substitution always holds as with complete information. On the contrary, if $\hat{\alpha}_t^i$ is increasing in q_t^{I}, then there are two competing effects: cultural substitution through $\hat{q}_t^i$ and cultural complementarity through $\hat{\alpha}_t^i$. Therefore, there exists cultural complementarity if and only if the (positive) effect of q_t^{I} on $\hat{q}_t^i$ is not too large with respect to its (positive) effect on $\hat{\alpha}_t^i$.

Proposition 4 regards the cultural complementarity or substitution displayed by each parent's socialization effort. However, the implications in terms of dynamics and stability crucially depend on how conjectures are distributed in the population, as shown in Proposition 3. Indeed, the population dynamics, as clear from equation (7), depends on the average socialization efforts, τ_t^{I} and τ_t^{J}. For example, let us consider stability at $q^{\mathrm{I}} = 0$. Even though all agents in both groups show cultural substitution, if agents I strongly overestimate their population shares (i.e., condition $S.0$ holds), then I's average effort is much lower than J's, and then $q^{\mathrm{I}} = 0$ can be stable.

Note that, in models with probabilistic transmission and complete information, cultural substitution is a sufficient condition to have cultural heterogeneity in the long run (Bisin and Verdier, 2001; Cheung and Wu, 2018). Previous propositions show that under incomplete information this may not hold, as we highlight in the following remark.

Remark. *Incomplete information may result in the coexistence of individual cultural substitution and long-run cultural homogeneity, or of individual cultural complementarity and long-run cultural heterogeneity.*

In Proposition 4, we have seen that a positive relationship existing between $\hat{\alpha}_t^i$ and q_t^{I} is a necessary but not sufficient condition for cultural complementarity. There may be different reasons and mechanisms whereby conjectured parental efficacy relates to population shares and, thus, the socialization effort displays cultural complementarity. A classical example is that proposed in Bisin and Verdier (2001), in which parental efficacy is assumed to be positively related to population shares and, in particular, $\alpha^i = q_t^{\mathrm{I}}$; incomplete information in that case may still lead to cultural complementarity whenever $\hat{q}_t^i$ is increasing in q_t^{I}, and i is aware of the functional form of α^i.

In the next section we show that, under incomplete information, if parents try to gain information during their interaction with children, a dependence of α^i on q_t^{I} may endogenously arise.

3. Conjectures' formation

We now endogenize the mechanism by which each parent $i \in I$ forms conjectures μ_t^i that induce the socialization effort τ_t^i.

Parent-child interaction develops over a long period of time, during which parents deal with incomplete information and shape their conjectures. To model this process, we assume that the parent-child interaction is composed by a period of **active learning** — i.e., effort experimentation and conjectures' updating — and an *actual* **socialization period**, during which a constant socialization effort is exerted and the child acquires a trait. In detail, the active learning period is composed of one or more subperiods. In each of these, each parent has the chance to experiment on their socialization effort. Given this experimentation, the parent receives some feedback and then updates their conjectures about the unknowns. At the end of each experimentation subperiod, each parent can decide either (i) to use the updated conjectures to produce the socialization effort and transmit their own trait, or (ii) to use the updated conjectures to produce a new experimentation effort so that a new experimentation subperiod starts. Crucially, we assume that the probability of acquiring the parent's trait depends only on the effort continuously exerted over the socialization period and it is unaffected by the efforts exerted during the experimentation subperiods.

In Section 3.1 we show what happens when parents exert just one experimentation effort, whereas in Section 3.2 we consider the case in which parents keep experimenting till the updating process converges. In both sections, we first characterize the locus of conjectures compatible with the updating process, then we show the optimal socialization efforts and discuss their properties.

3.1. Single experimentation effort

Each parent $i \in I$ of cohort t, at experimentation subperiod 0, starts with conjectures $\mu^i_{t,0} = (\mu^i_{\alpha^i,t,0}, \mu^i_{q^{\mathrm{I}}_t,t,0}) \in \Delta(\mathbb{R}_+) \times [0,1]$. To simplify the notation, in what follows we drop the index t because we are analyzing the learning process within each cohort. Given the conjectures μ^i_0, each parent i exerts an **experimentation effort** $\tilde{\tau}^i_0 := \hat{\alpha}^i_0(1-\hat{q}^i_0)\Delta V^{\mathrm{I}}$ and, then, receives a series of signals $(s^i_{k,0})^{\mathrm{K}}_{k=1} \in [0,1]^{\mathrm{K}}$ that are messages the child delivers about the likelihood they have to acquire parent's trait if the socialization effort τ^i were to coincide with the experimentation effort $\tilde{\tau}^i_0$ itself. Consistent with this, we assume signals to have mean $p^{i\mathrm{I}}(\tilde{\tau}^i_0)$, which is the transition probability given $\tilde{\tau}^i_0$. We further assume that signals are generated by a probability distribution belonging to the exponential family and that each parent starts with a corresponding conjugate prior.

Denote by $\pi^i_0 \in \Delta([0,1])$ the parent i's *prior* about the transition probability. Thus, i's prior expectation about transition probability, given the experimentation effort $\tilde{\tau}^i_0$, is $\hat{p}^{i\mathrm{I}}_0(\tilde{\tau}^i_0) := \mathbb{E}^i_{\pi^i_0}[p^{i\mathrm{I}}(\tilde{\tau}^i_0)] = \hat{q}^i_0 + \hat{\alpha}^i_0\tilde{\tau}^i_0(1-\hat{q}^i_0)$.

Given the prior π^i_0 and the sequence of signals $(s^i_{k,0})^{\mathrm{K}}_{k=1}$ induced by the experimentation effort $\tilde{\tau}^i_0$, each parent i uses Bayesian inference to form a *posterior* $\pi^i_1 \in \Delta([0,1])$ and sets new conjectures μ^i_1 compatible with the posterior π^i_1.[14] The mean of the posterior is defined as $\hat{p}^{i\mathrm{I}}_1(\tilde{\tau}^i_0) := \mathbb{E}^i_{\pi^i_1}[p^{i\mathrm{I}}(\tilde{\tau}^i_0)] = \hat{q}^i_1 + \hat{\alpha}^i_1\tilde{\tau}^i_0(1-\hat{q}^i_1)$, where $(\hat{\alpha}^i_1, \hat{q}^i_1)$ are the means of the updated conjectures about the unknowns.

As we can see in equation (4), the expectations of unknowns are the unique determinants of the socialization effort. The following proposition characterizes the locus of $(\hat{\alpha}^i_1, \hat{q}^i_1)$ compatible with the Bayesian updating, given the experimentation effort $\tilde{\tau}^i_0$. Define the sample mean of signals as $\bar{s}^i_{\mathrm{K},0} := \frac{\sum^{\mathrm{K}}_{k=1} s^i_{k,0}}{\mathrm{K}}$.

[14] Note that the parent can do this because they know the functional form of the transition probability and they recall their own experimentation effort.

Proposition 5. *For each $I \in \mathcal{C}$ and $i \in I$, given μ_0^i and a number κ of signals induced by $\tilde{\tau}_0^i$, there exists a $\lambda_{\kappa,0}^i \in [0,1]$ such that, given the updated conjectures μ_1^i, the pair $(\hat{\alpha}_1^i, \hat{q}_1^i)$ satisfies the following:*

$$\underbrace{\hat{p}_1^{i\mathrm{I}}(\tilde{\tau}_0^i)}_{posterior\ mean} = \lambda_{\kappa,0}^i \cdot \underbrace{\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i)}_{prior\ mean} + (1-\lambda_{\kappa,0}^i) \cdot \underbrace{\bar{s}_{\kappa,0}^i}_{sample\ mean}, \tag{9}$$

with $\lim_{\kappa\to\infty} \lambda_{\kappa,0}^i = 0$ *and* $\lim_{\kappa\to\infty} \bar{s}_{\kappa,0}^i = p^{i\mathrm{I}}(\tilde{\tau}_0)$.

From (9), if $\bar{s}_{\kappa,0}^i = p^{i\mathrm{I}}(\tilde{\tau}_0)$, the (posterior) bias of each agents i, $b_1^i := \hat{q}_1^i - q^{\mathrm{I}}$, satisfies

$$b_1^i = \frac{-(\hat{\alpha}_1^i - \alpha^i)\tilde{\tau}_0^i(1-q^{\mathrm{I}}) + \lambda_{\kappa,0}^i\Big(\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^i)\Big)}{(1-\hat{\alpha}_1^i\tilde{\tau}_0^i)}, \tag{10}$$

where $\hat{q}_1^i < \hat{p}_1^{i\mathrm{I}}(\tilde{\tau}_0^i)$.

Equation (9) states that the posterior mean, $\hat{p}_1^{i\mathrm{I}}(\tilde{\tau}_0^i)$, is a convex linear combination between the prior mean, $\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i)$, and the mean of the signals received, $\bar{s}_{k,0}^i$. The weighting $\lambda_{\kappa,0}^i$ depends on the number signals. The higher κ the lower $\lambda_{\kappa,0}^i$, meaning that, forming the posterior, each agent is less affected by the value of the prior and more by the mean of the signals, which, in turn, becomes closer to the actual mean. Thus, if the number of signals is small, agents assign small weight to the empirical observation with respect to their prior, whereas if the number of signals increases, agents weight the mean of signals more. At the limit, if the agent receives an infinite number of signals, then the mean of the signals always coincides with the theoretical transition probability given the experimentation effort, $p^{i\mathrm{I}}(\tilde{\tau}_0^i)$, and, thus, the posterior mean $\hat{p}_1^{i\mathrm{I}}(\tilde{\tau}_0^i)$ coincides with it. Notably, recalling that $\hat{p}_1^{i\mathrm{I}}(\tilde{\tau}_0^i) = \hat{q}_1^i + \hat{\alpha}_1^i\tilde{\tau}_0^i(1-\hat{q}_1^i)$, there might be infinitely many $(\hat{\alpha}_1^i, \hat{q}_1^i)$ compatible with the posterior mean. As a consequence, parents, using the information gained through one experimentation effort, generally fail to learn the true parameters.

The second part of the proposition analyzes the case in which the mean of signals is correct, i.e., $\bar{s}_{\kappa,0}^i = p^{i\mathrm{I}}(\tilde{\tau}_0^i)$. We use it as a maintained assumption in what follows. This can happen either for large κ or if the signals are extremely precise (e.g., children are able to communicate precisely the probability of their acquiring a trait). Equation (10) provides a reformulation of the locus $(\hat{\alpha}_1^i, \hat{q}_1^i)$ compatible with the Bayesian update. From this, we can see how the posterior bias on population share's conjecture, b_1^i, negatively relates to the posterior overestimation of parental efficacy, $(\hat{\alpha}_1^i - \alpha^i)$. Indeed, looking at the RHS and LHS of equation (10), there exists an inverse relationship between $\hat{q}_1^i$ and $\hat{\alpha}_1^i$, so that, if the bias about population share's conjecture is positive (i.e., $b_1^i > 0$), then agent i underestimates own parental efficacy (i.e., $\hat{\alpha}_1^i < \alpha^i$).

We can also see the role played by the number of signals, κ. If agent i receives an infinite number of signals, then $b_1^i = -\frac{(\hat{\alpha}_1^i - \alpha^i)\tilde{\tau}_0^i(1-q^{\mathrm{I}})}{(1-\hat{\alpha}_1^i\tilde{\tau}_0^i)}$, which implies that correct mean conjecture about α^i implies correct conjecture about q^{I}, and vice versa. On the contrary, if κ is finite (which implies $\lambda_{\kappa,0}^i > 0$), even if agent i has correct mean conjecture about one unknown, the limited number of signals does not allow them to learn the other unknown (e.g., if $\hat{\alpha}^i = \alpha^i$, then $b_1^i = \frac{\lambda_{\kappa,0}^i(\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^i))}{(1-\hat{\alpha}_1^i\tilde{\tau}_0^i)}$).

Lastly, Proposition 5 also provides an upper-bound for the conjectures about population shares, namely $\hat{q}_1^i < \hat{p}_1^{i\mathrm{I}}(\tilde{\tau}_0^i)$. This must always hold because the posterior transition probability is a convex linear combination between 1 and $\hat{q}_1^i$.

Optimal socialization effort The optimal socialization effort after a single experimentation is given by equation (4) evaluated at a pair $(\hat{\alpha}_1^i, \hat{q}_1^i)$ satisfying locus in equation (10). We assume that parent i derives $\hat{\alpha}_1^i$ from equation (10) for any given $\hat{q}_1^i$.[15] Thus, given any $\hat{q}_1^i < \hat{p}_1^{i\mathrm{I}}(\tilde{\tau}_0^i)$, the mean conjecture about parental efficacy of parent i that satisfies (10) is such that

$$\hat{\alpha}_1^i = \frac{\alpha^i(1-q^{\mathrm{I}})}{(1-\hat{q}_1^i)} + \frac{\lambda_{\mathrm{K},0}^i\big(\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^i)\big) - b_1^i}{\tilde{\tau}_0^i(1-\hat{q}_1^i)}. \tag{11}$$

Therefore, each parent i who undertakes just one experimentation step, uses the updated conjectures to exert the socialization effort, so that $\hat{\alpha}^i \equiv \hat{\alpha}_1^i$, $\hat{q}^i \equiv \hat{q}_1^i$. Thus $\tau^i \equiv \tilde{\tau}_1^i$ and it takes the following form:

$$\tau^i = \tau_{bv}^i + \left[\lambda_{\mathrm{K},0}^i\Big(\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^i)\Big) - b^i\right]\frac{\Delta V^{\mathrm{I}}}{\tilde{\tau}_0^i}. \tag{12}$$

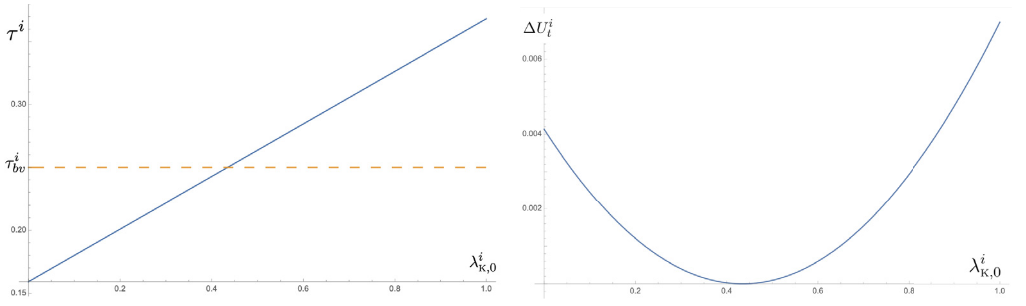
From equation (12), we can see how τ^i can be written as the socialization effort under complete information, τ_{bv}^i, plus a term internalizing the distortion with respect to the objectively optimal socialization effort. We can clearly see that the distortion is increasing in the cultural intolerance, ΔV^{I}, and decreasing in the experimentation effort, $\tilde{\tau}_0^i$. The increasing effect of ΔV^{I} is due to the fact, as it can be seen in equation (4), that the intolerance magnifies the effect of wrong conjectures on τ^i. Instead, concerning the decreasing effect of $\tilde{\tau}_0^i$, the lower the effort the closer the transition probability $p^{i\mathrm{I}}(\tilde{\tau}_0^i)$ is to the population share and, thus, the effect of wrong conjectures about q^{I} is more distortive.

From equation (12), we can also see that the larger the number of signals is, the more the sign and magnitude of the distortion depend on the bias, $b^i = \hat{q}^i - q^{\mathrm{I}}$. Indeed, when parent i receives infinite signals, $\lambda_{\mathrm{K},0}^i = 0$ and $\tau^i = \tau_{bv}^i - b^i\frac{\Delta V^{\mathrm{I}}}{\tilde{\tau}_0^i}$. On the contrary, for a finite number of signals, $\lambda_{\mathrm{K},0}^i > 0$ and the distortion is affected also by the difference between conjectured and *actual* transition probabilities, so that τ^i can be closer to τ_{bv}^i than it is when parents receive an infinite number of signals.

Note that the difference between subjective and objective optimal socialization efforts also determines the magnitude of individual utility loss shown in Proposition 2. In Fig. 2a, we can see a numerical example showing how τ^i depends on $\lambda_{\mathrm{K},0}^i$ (which, in turn, negatively depends on K), compared to τ_{bv}^i (which is independent of K). Fig. 2b shows the individual utility loss associated with τ^i. Notably, increasing the number of signals does not always result in a lower distortion or in a lower individual utility loss.

We now discuss the results of comparative statics as in Proposition 4, considering the subjectively optimal socialization effort τ^i as characterized in equation (12). For this purpose, we consider two examples to see how cultural substitution or complementarity is displayed as dependent of different assumptions about how $\hat{q}^i$ depends on q_t^{I} and how, in turn, it affects $\hat{\alpha}^i$

[15] This mimics the fact that agents first form conjectures about population shares, which are more susceptible to being affected by external sources (as we do in Section 4 with cultural leaders), and then form conjectures about parental efficacy, which should stem from a process of self-assessment.



(a) (b)

Fig. 2. Effect of $\lambda^i_{\mathrm{K},0}$ on (a) subjectively optimal socialization effort, τ^i, when a parent exerts a single experimentation effort, and (b) individual utility loss given τ^i. Parameterization: $q^{\mathrm{I}} = \alpha^i = \frac{1}{2}$, $\hat{q}^i = 0.45$, $\Delta V^{\mathrm{I}} = 1$.

according to equation (11). In the first example, we show what happens when $\hat{q}^i$ does not react to changes in q^{I}, whereas the second example considers the case in which the conjecture about population share linearly depends on q^{I}.

Example 1. Let us consider $\hat{q}^i$ to be independent of q^I. In this case, any shock about population shares is fully internalized by parent i in their updated parental efficacy, according to equation (11). We always observe cultural complementarity. Indeed,

$$\frac{\partial \tau^i}{\partial q^{\mathrm{I}}} = \frac{(1-\alpha^i \tilde{\tau}^i_0)(1-\lambda^i_{\mathrm{K},0})}{\tilde{\tau}^i_0} \Delta V^i > 0.$$

This result stems from the fact that τ^i is increasing in $\hat{\alpha}^i$ which, in turn, is increasing in q^I, as shown in (4).

Example 2. Let us consider $\hat{q}^i$ to be linearly dependent on q^I with a fixed bias $\tilde{b}^i \in [-q^I, 1-q^I]$. In detail, we assume that $\hat{q}^i \equiv \hat{q}^i_1 = \hat{q}^i_0 = q^I + \tilde{b}^i$. In this case, shocks about population shares affect both $\hat{q}^i$ (directly) and $\hat{\alpha}^i$ (indirectly through the update in (11)). Differentiating (12) with respect to q^I, we can easily see that: (i) if $\tilde{b}^i < 0$, there is always cultural substitution; (ii) if $\tilde{b}^i > 0$ there is cultural complementarity if and only if $\alpha^i < \frac{\tilde{b}^i}{\tilde{\tau}^i_0(1-q^{\mathrm{I}}-\tilde{b}^i)} - \frac{\lambda^i_{\mathrm{K},0}}{1-\lambda^i_{\mathrm{K},0}}\hat{\alpha}^i_0$.

The difference between the two examples above is that, even if in both of them agents have incomplete information about q^{I}, in Example 2 agents correctly increase/decrease the conjecture, responding to changes in population shares. Therefore, in such a case, parents have an additional piece of information that makes Example 2 closer to the case of complete information. This explains why in Example 2 we have a parameters space such that cultural substitution is shown (as it always happens with complete information), whereas in Example 1 there is only cultural complementarity.

We can also compare these two examples with the general result of comparative statics in Section 2.2. The result of Example 1 trivially follows from Proposition 4. Indeed, the condition for cultural complementarity is $\frac{\partial \hat{q}^i}{\partial q^{\mathrm{I}}} < \frac{(1-\hat{q}^i)}{\hat{\alpha}^i}\frac{\partial \hat{\alpha}^i}{\partial q^{\mathrm{I}}}$, which now reads $0 < \frac{(1-\hat{q}^i)}{\hat{\alpha}^i} \cdot \frac{(1-(1-\hat{q}^i_0)\alpha^i \hat{\alpha}^i_0 \Delta V^{\mathrm{I}})(1-\lambda^i_{\mathrm{K},0})}{(1-\hat{q}^i_0)(1-\hat{q}^i_1)\Delta\hat{\alpha}^i_0 V^{\mathrm{I}}}$ which is always satisfied. On the other hand, the fact that, in Example 2, for $\tilde{b}^i < 0$ there is always cultural substitution can also be seen from the general result in Proposition 4, which states that whenever $\frac{\partial \hat{\alpha}^i}{\partial q^{\mathrm{I}}} < 0$ there is always cultural substitution. Indeed, if $\tilde{b}^i < 0$, we have that $\frac{\partial \hat{\alpha}^i}{\partial q^{\mathrm{I}}} = \tilde{b}^i \cdot \frac{(2-(1-\hat{q}^i)\Delta V^{\mathrm{I}} \alpha^i \hat{\alpha}_0^i)(1-\lambda_{\mathrm{K},0})}{(1-\hat{q}^i)^3 \Delta V^{\mathrm{I}} \hat{\alpha}_0} < 0$.

3.2. Converged active learning

Considering the analysis of the previous section, with a single experimentation effort we generally have that $\tilde{\tau}_1^i \neq \tilde{\tau}_0^i$, i.e., the effort parent i would exert with conjectures μ_1^i is different from the experimentation effort they used to obtain those conjectures. Thus, there is room for parents to keep experimenting to gain additional information from further experimentation steps. In this section, we consider the case in which parents keep experimenting till the updating process converges. The main result we find is that, even in this case, parents may still fail to learn the unknowns and, thus, they generally choose a suboptimal socialization effort.

Consider many experimentation efforts such that, at each experimentation subperiod $n \in \mathbb{N}_0$, each parent has prior conjectures μ_n^i, produces a corresponding experimentation effort $\tilde{\tau}_n^i$, and after receiving a series of signals $(s_{k,n}^i)_{k=1}^{\mathrm{K}}$, they form their updated conjectures μ_{n+1}^i based on which they exert a new experimentation effort $\tilde{\tau}_{n+1}^i$. In this respect, for a generic n, keeping the assumption $\bar{s}_{\mathrm{K},n}^i = p^{i\mathrm{I}}(\tilde{\tau}_n^i)$ and given the updated conjectures μ_{t+1}^i, the pair $(\hat{\alpha}_{t+1}^i, \hat{q}_{t+1}^i)$ must satisfy

$$\hat{p}_{n+1}^{i\mathrm{I}}(\tilde{\tau}_n^i) = \lambda_{\mathrm{K},n}^i \cdot \hat{p}_n^{i\mathrm{I}}(\tilde{\tau}_n^i) + (1-\lambda_{\mathrm{K},n}^i) \cdot p^{i\mathrm{I}}(\tilde{\tau}_n^i). \tag{13}$$

Note that equation (13) is the counterpart of equation (9) for a generic n. If the process converges, at steady state $\hat{p}^{i\mathrm{I}}(\tilde{\tau}_n^i) \equiv \hat{p}_{n+1}^{i\mathrm{I}}(\tilde{\tau}_n^i) = \hat{p}_n^{i\mathrm{I}}(\tilde{\tau}_n^i)$. That is, given the experimentation effort exerted $\tilde{\tau}_n^i$, the posterior conjectured transition probability is equal to the prior conjectured transition probability. We further assume that, at steady state, parent i keeps conjectures unchanged among all the possible pairs compatible with (13). Thus, $(\hat{\alpha}^i, \hat{q}^i) \equiv (\hat{\alpha}_{n+1}^i, \hat{q}_{n+1}^i) = (\hat{\alpha}_n^i, \hat{q}_n^i)$ and, $\tau^i \equiv \tilde{\tau}_n^i = \tilde{\tau}_{n-1}^i$.

Proposition 6. *For each $I \in \mathcal{C}$ and $i \in I$, given any μ_0^i, at steady state of equation (13), $\hat{p}^{i\mathrm{I}}(\tau^i) = p^{i\mathrm{I}}(\tau^i)$. Therefore, the pair $(\hat{\alpha}^i, \hat{q}^i)$ satisfies the following*

$$b^i = \left[(\alpha^i(1-q^{\mathrm{I}}) - \hat{\alpha}^i(1-\hat{q}^i)\right]\hat{\alpha}^i(1-\hat{q}^i)\Delta V^{\mathrm{I}}. \tag{14}$$

Moreover, $b^i < \bar{b}^i := \frac{1}{4}\Delta V^{\mathrm{I}}[\alpha^i(1-q^{\mathrm{I}})]^2$. A sufficient condition for equation (13) to converge is that $\mathrm{K} \to \infty$.

Equation (14) displays the locus of points (10) when the active learning process has converged. Note that, as we have argued above, although when reaching the steady state parent i correctly conjectures the mean transition probability $p^{i\mathrm{I}}(\tau^i)$, this may not be sufficient for a correct inference about the underlying parameters $(\alpha^i, q^{\mathrm{I}})$, and, according to (14), many pairs $(\hat{\alpha}^i, \hat{q}^i)$ are compatible with the converged Bayesian learning. Notably, we can verify from equation (14) that, if parent i has knowledge of either α^i or q^{I}, then they can correctly infer the other unknown parameter.

Note that the steady state condition $\hat{p}^{i\mathrm{I}}(\tau^i) = p^{i\mathrm{I}}(\tau^i)$ implies that the conjectures that induce the socialization effort are those that are *confirmed* by the experimentation. Therefore, *subjectively optimal* socialization effort in equation (4), given pairs $(\hat{\alpha}^i, \hat{q}^i)$ satisfying condition

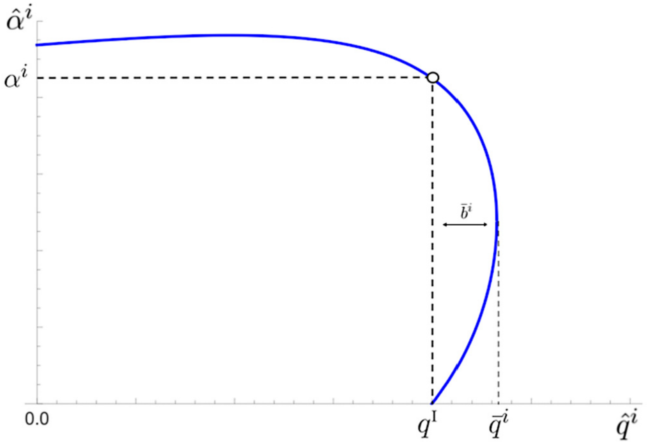


Fig. 3. Representation of locus $(\hat{\alpha}^i, \hat{q}^i)$ satisfying equation (14).

(14), also constitutes a **selfconfirming equilibrium** of the parental socialization problem with transition probability as feedback, which requires (i) subjective rationality and (ii) confirmed conjectures (refer to Battigalli, 1987; Fudenberg and Levine, 1993; Battigalli et al., 2015).[16]

To better understand the implications of the locus in equation (10), consider now Fig. 3. We can see that $\hat{\alpha}^i > \alpha^i$ if and only if $\hat{q}^i < q^{\mathrm{I}}$ (i.e., $b^i < 0$), namely, at steady state, parents underestimate their parental efficacy if and only if they overestimate the share of their own cultural group. Note also that: (i) $\bar{b}^i$ characterizes the magnitude of the maximum bias compatible with equation (14) and, thus, if $\hat{q}^i > \bar{q}^i := q^{\mathrm{I}} + \bar{b}^i$, no $\hat{\alpha}^i$ can be found; (ii) if $q^{\mathrm{I}} < \hat{q}^i < \bar{q}^i$, two $\hat{\alpha}^i$'s exist (α_h^i and α_l^i); and (iii) if $\hat{q}^i < q^{\mathrm{I}}$, only $\hat{\alpha}_l^i$ exists. By means of simple algebra we get:

$$\hat{\alpha}_h^i := \frac{\tau_{bv}^i + \xi(b^i)}{2(1-\hat{q}^i)\Delta V^{\mathrm{I}}}, \qquad \hat{\alpha}_l^i := \frac{\tau_{bv}^i - \xi(b^i)}{2(1-\hat{q}^i)\Delta V^{\mathrm{I}}}, \tag{15}$$

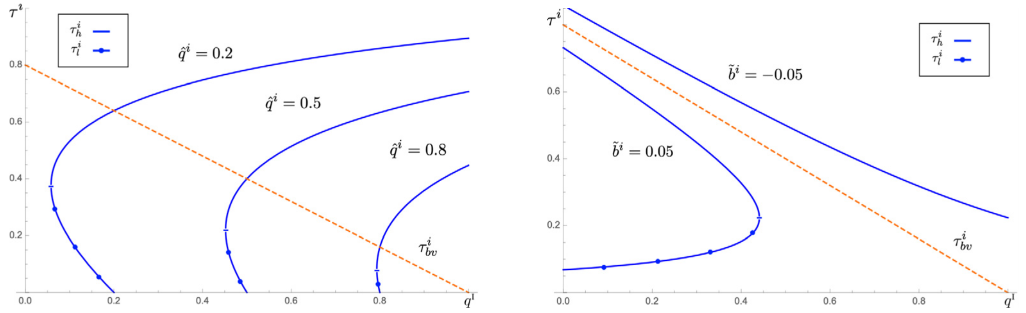
where $\xi(b^i) := 2\sqrt{\Delta V^{\mathrm{I}}\left(\bar{b}^i - b^i\right)}$. Therefore, for each $i \in I$ and each $\hat{q}^i < \bar{q}^i$, two possible optimal socialization efforts associated with $\hat{\alpha}_h^i$ and $\hat{\alpha}_l^i$ exist:

$$\tau_h^i := \frac{\tau_{bv}^i + \xi(b^i)}{2}, \quad \tau_l^i := \frac{\tau_{bv}^i - \xi(b^i)}{2}. \tag{16}$$

Note that if $b^i = \bar{b}^i$ (i.e., $\hat{q}^i = \bar{q}^i$) then $\tau_h^i = \tau_l^i = \frac{1}{2}\tau_{bv}^i$. From equation (16), any socialization effort can be expressed as the algebraic mean between the objectively optimal effort, τ_{bv}^i, and a function of the agent's bias about population share, b^i.

Consider now the comparative statics for the socialization efforts. As we did for parents exerting a single experimentation effort, we consider the case in which $\hat{q}^i$ is fixed, i.e., it does not

[16] This analogy is not surprising since it is an established result that any selfconfirming equilibrium is a steady state of a generic adaptive learning, if it converges (see, for example, Battigalli et al., 1992; Fudenberg and Kreps, 1995; Gilli, 1999; Battigalli et al., 2019, 2021). Moreover, note that the parental experimentation problem can be thought of as a stochastic control problem and Battigalli and Lanzani (2020) relates the selfconfirming equilibrium concept in decision problems with feedback (e.g., Battigalli et al., 2015) to the limit behavior of solutions to stochastic control problems (e.g., Easley and Kiefer, 1988). For this reason, our result in Proposition 6 holds also relaxing the assumption that $\bar{s}_{\mathrm{K},n}^i = p^{i\mathrm{I}}(\tilde{\tau}_n^i)$.



(a) exogenous $\hat{q}^i$ (as in Example 1) (b) $\hat{q}^i = q^{\mathrm{I}} + \tilde{b}^i$ (as in Example 2)

Fig. 4. Optimal socialization efforts given different assumptions about $\hat{q}^i$ and $\alpha^i = 0.8$, $\Delta V^{\mathrm{I}} = 1$.

move together with q^{I} (as in Example 1), and the case in which agents keep the bias b^i fixed as q^{I} changes (as in Example 2). In the case of Example 1 (represented in Fig. 4a), τ_h^i always displays cultural complementarity, whereas τ_l^i always displays cultural substitution. On the contrary, in the case of Example 2 (Fig. 4b), τ_h^i always displays cultural substitution, whereas τ_l^i, when it exists ($\tilde{b}^i > 0$), always displays cultural complementarity.

Having discussed the properties of the subjectively optimal effort at the steady state, we conclude by showing in Fig. 5 a numerical example of how τ^i changes together with the number n of parental experimentation, and how it relates to the objectively optimal effort, τ_{bv}^i. Notably, increasing the number of experimentation does not ensure that parent chooses an effort closer to the objectively optimal one. For the numerical example in Fig. 5 we chose $\hat{q}^i < q^{\mathrm{I}}$, which implies that the subjectively optimal socialization effort converges to τ_h^i. Had we chosen $\hat{q}^i > q^{\mathrm{I}}$, we would have observed effort converging to either τ_h^i or τ_l^i, depending on the starting conditions.

4. Cultural leaders: a discussion

Incomplete information about population shares is a particularly relevant issue because conjectures about q_t^{I} can be shaped by social media and (fake) news, and can be manipulated and exploited by cultural or political leaders.

Let us introduce, for each group $I \in \mathcal{C}$, a cultural leader ℓ^{I} who knows q_t^{I} and who can instill in the agents of their own group possibly biased conjectures. For the sake of simplicity, in this discussion, we assume that agents are homogeneous within groups, so that the leader ℓ^{I} instills $\hat{q}_t^{\mathrm{I}}$ in the representative agent.[17] The leader ℓ^{I} chooses an action $a_t^{\ell^{\mathrm{I}}} \in A^{\ell^{\mathrm{I}}} := \{\hat{q}_t^{\mathrm{I}+}, q_t^{\mathrm{I}}, \hat{q}_t^{\mathrm{I}-}\}$, where $\hat{q}_t^{\mathrm{I}+}$ stands for conjectures with an exogenously given positive bias, $\hat{q}_t^{\mathrm{I}-}$ for conjectures with an exogenously given negative bias, and q_t^{I} for correct conjectures (i.e., a null bias).[18]

At each t, we consider two possible types of leader: an *identitarian* leader, who maximizes the share of their own cultural group in the next generation, and a *utilitarian* leader, who cares about the realized utility of the representative agent of own group. Therefore, for each group $I \in \mathcal{C}$, each leader faces the following problem:

[17] Note that our results are unchanged even when considering heterogeneous agents and individual specific conjectures $\hat{q}_t^i$, as in the other sections of the paper.

[18] To simplify the discussion, note that we consider leaders who can choose only the sign, and not the magnitude of the bias.

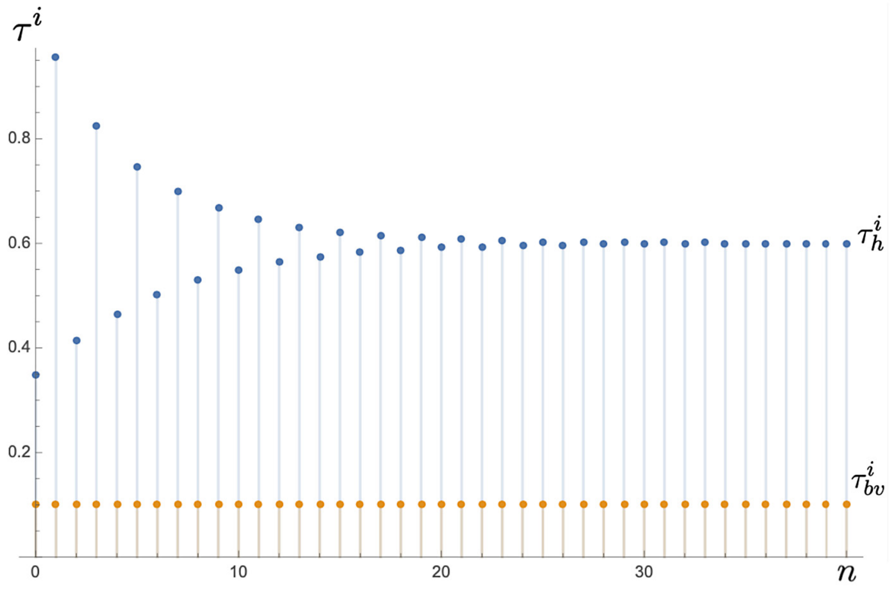


Fig. 5. Socialization effort τ^i depending on the number of experimentation, n. Parameterization $\Delta V^{\mathrm{I}} = 1$, $\alpha^i = \frac{1}{2}$, $q^{\mathrm{I}} = \frac{4}{5}$, and, for each n, $\hat{q}^i_n = \frac{1}{2}$.

$$\max_{a_t^{\ell^{\mathrm{I}}} \in A^{\ell^{\mathrm{I}}}} u^{\ell^{\mathrm{I}}}\left(a_t^{\ell^{\mathrm{I}}}; a_t^{\ell^{\mathrm{J}}}\right) = \begin{cases} q^{\mathrm{I}}_{t+1}\left(a_t^{\ell^{\mathrm{I}}}; a_t^{\ell^{\mathrm{J}}}\right) & \text{if identitarian,} \\ u^{\mathrm{I}}\left(a_t^{\ell^{\mathrm{I}}}; a_t^{\ell^{\mathrm{J}}}\right) & \text{if utilitarian.} \end{cases} \tag{17}$$

Let us now discuss the optimal action of the leader depending on how the representative agent forms their conjectures. In such a case, the leader should take into account the fact that I's conjectured parental efficacy depends on the conjecture induced by the leader themself, as described in (11) for the case of single experimentation effort (for this case we consider only $\mathrm{K} \to \infty$),[19] and in (15) for the case of converged active learning.

A utilitarian leader always instills a correct conjecture about population shares in the members of own group, $a_t^{\ell^{\mathrm{I}}} = q_t^{\mathrm{I}}$. This because, as shown in Section 3, correct conjectures about q^{I} induce correct conjectures about parental efficacies. This, together with Proposition 2, in which we show that the maximum utility for each agent is reached when parents have correct conjectures, explains the utilitarian leader's choice. On the other hand, an identitarian leader always chooses to instill a negative bias, $a_t^{\ell^{\mathrm{I}}} = \hat{q}_t^{\mathrm{I}-}$. This result derives from the fact that the choice of a negative bias, compared to a positive one, always induces a higher socialization effort and an increase in the share of own cultural group in the next generation.

We can now discuss the impact of cultural leaders on long-run outcomes. Recall that Proposition 3 and Fig. 1 show how the long-run configurations depend only on agents' biases about population shares. Let us consider the leaders' choices as discussed above. If the two leaders are utilitarian they instill correct conjectures, so that the dynamics is equivalent to that with complete information and it is qualitatively represented by Fig. 1c. On the other hand, if at least one leader is identitarian, they can instill negative biases and, according to Proposition 3, a homogeneous

[19] Results also hold unchanged for a finite K if $\lambda^i_{\mathrm{K},0}$ is small enough.

population may be the long-run result of the dynamics (Figs. 1a-b-d). The specific long-run outcome depends on the strength of the biases. Next, we provide an example clarifying this point.

Example 3. Let us consider parents who endogenously choose the socialization effort after the process of active learning has converged (as in Section 3.2) and, for each $I \in \mathcal{C}$, the conjecture induced by the cultural leader takes the form $\hat{q}_t^{\mathrm{I}} = (1 - \beta^{\mathrm{I}})q_t^I$, with $\beta^I \in [0, 1]$. Note that the optimal choice of an identitarian leader represented by $\beta^I > 0$, whereas the choice of a utilitarian one is represented by $\beta^I = 0$.

Applying results of Proposition 3 to this example, we can see that:

S.0 $q^I = 0$ is stable if and only if $\beta^J > \bar{\beta}^{\mathrm{I}} := \frac{\alpha^{\mathrm{J}4}}{\alpha^{\mathrm{I}2}} \frac{\Delta V^{\mathrm{J}2}}{\Delta V^{\mathrm{I}}}$,
S.1 $q^I = 1$ is stable if and only if $\beta^I > \bar{\beta}^{\mathrm{J}} := \frac{\alpha^{\mathrm{I}4}}{\alpha^{\mathrm{J}2}} \frac{\Delta V^{\mathrm{I}2}}{\Delta V^{\mathrm{J}}}$.

Note also that since, for each $I \in \mathcal{C}$, $\hat{q}^I$ is linear in q^I, it satisfies the conditions of Proposition 3 for having a unique interior steady state, $q^* \in (0, 1)$. Therefore, q^* is globally stable if and only if $\beta^I < \bar{\beta}^{\mathrm{I}}$ and $\beta^J < \bar{\beta}^{\mathrm{J}}$.

In detail, focusing on the stability of $q^I = 0$, the more I is effective in the vertical transmission, α^I, and the greater their intolerance, ΔV^I, the more a large β^J is needed for J agents to invade the society and have q^I stable; α^J and ΔV^J have the opposite effect.

Given the parameters $(\alpha^I, \Delta V^I)_{I \in \mathcal{C}}$, the cultural dynamics can show all the cases of Fig. 1, namely, both stable and unstable polymorphic equilibria, and also no polymorphic equilibria at all. In detail, if both leaders do not instill too negative a bias ($\beta^I < \bar{\beta}^{\mathrm{I}}$ and $\beta^J < \bar{\beta}^{\mathrm{J}}$), the efforts are close to the complete information case, and the standard result of stable polymorphic equilibrium holds (Fig. 1c). On the contrary, if at least one leader can induce strong negative bias, we always observe long-run cultural homogeneity. If $\beta^I > \bar{\beta}^{\mathrm{I}}$ and $\beta^J > \bar{\beta}^{\mathrm{J}}$, then there exists a unique polymorphic equilibrium which is also unstable (Fig. 1b), whereas if just one leader can instill strong negative bias, either $q^I = 0$ (Fig. 1a) or $q^I = 1$ (Fig. 1d) is globally stable.

5. Conclusion

In this paper, we propose a generalization of the standard cultural transmission framework (Bisin and Verdier, 2001) to the case of incomplete information. In detail, we introduce incomplete information in the two key dimensions on which cultural transmission operates: vertical socialization — through parental efficacy — and oblique socialization — through population shares. We find that the long-run composition of the society mainly depends on the biases in conjectures about population shares. Moreover, depending on the properties of conjectures about parental efficacy, optimal socialization efforts may also display cultural complementarity; we also show that, under incomplete information and differently than under complete information, cultural substitution is not a sufficient condition to achieve long-run cultural heterogeneity. Lastly, in our framework, the two dimensions of incomplete information jointly determine the socialization outcome so that, even if parents try to learn through experimentation, they generally fail to do so.

This paper is a first attempt at introducing incomplete information in cultural transmission and, for this purpose, we employed the baseline framework. Consequently, there are many important dimensions that we did not consider and that deserve further investigation. For example, one of the first sources of incomplete information is the ignorance of parents about the specific friendship networks children are exposed to. Moreover, an important dimension to consider is that children may have an active role in the socialization process, and their choices are partially hidden from parents. This issue relates also to the growing literature about the role of parenting styles affecting socialization and other economically relevant outcomes. Note also that, incomplete information is a key element in the analysis of many economic phenomena and, at the same time, it is important to study links between cultural dynamics and economic outcomes. Therefore, an interesting direction would be to integrate the uncertainty stemming from the economic environment with parents' socialization choices in the cultural transmission framework. From an applied perspective, the framework we propose is also strictly related to the present debate about the spread of fake news that agents are unable to detect as such. Lastly, a specific focus on leaders exploiting incomplete information to induce specific cultural dynamics could shed some light on the debate about populism.

Appendix A

Proof of Proposition 1. Recall that $p^{i\mathrm{I}}_{bv,t} = q^{\mathrm{I}}_t + \alpha^i_t \tau^i_{bv,t}(1-q^{\mathrm{I}}_t)$ and $p^{i\mathrm{I}}_t = q^{\mathrm{I}}_t + \alpha^i_t \tau^i_t (1-q^{\mathrm{I}}_t)$. Thus, given equation (4),

$$\begin{aligned}
p^{i\mathrm{I}}_t > p^{i\mathrm{I}}_{bv,t} \quad &\Leftrightarrow \quad q^{\mathrm{I}}_t + \alpha^i_t \underbrace{\hat{\alpha}^i_t(1-\hat{q}^i_t)\Delta V^{\mathrm{I}}}_{\tau^i}(1-q^{\mathrm{I}}_t) > q^{\mathrm{I}}_t + \alpha^i_t \underbrace{\alpha^i_t(1-q^{\mathrm{I}}_t)\Delta V^{\mathrm{I}}}_{\tau^i_{bv,t}}(1-q^{\mathrm{I}}_t) \\
&\Leftrightarrow \quad \hat{\alpha}^i_t(1-\hat{q}^i_t) > \alpha^i(1-q^{\mathrm{I}}_t) \\
&\Leftrightarrow \quad -\hat{q}^i_t > -1 + \frac{\alpha^i_t}{\hat{\alpha}^i_t}(1-q^{\mathrm{I}}_t) \\
&\Leftrightarrow \quad \hat{q}^i_t < 1 - \frac{\alpha^i_t}{\hat{\alpha}^i_t}(1-q^{\mathrm{I}}_t) \\
&\Leftrightarrow \quad \hat{q}^i_t - q^{\mathrm{I}}_t < 1 - q^{\mathrm{I}}_t - \frac{\alpha^i_t}{\hat{\alpha}^i_t}(1-q^{\mathrm{I}}_t) \\
&\Leftrightarrow \quad b^i_t < (1-q^{\mathrm{I}}_t)\left(1 - \frac{\alpha^i_t}{\hat{\alpha}^i_t}\right). \quad \square
\end{aligned}$$

Proof of Proposition 2. Let us recall the objective expected utility given subjective and objective optimal socialization effort

$$\mathbb{E}_{p^{i\mathrm{I}}}[u(\tau^i_t)] = V^{\mathrm{IJ}} + \Delta V^{\mathrm{I}}\left(q^{\mathrm{I}}_t + (1-q^{\mathrm{I}})\alpha^i \tau^i_t\right) - \frac{1}{2}\left(\tau^i_t\right)^2.$$

and

$$\mathbb{E}_{p^{i\mathrm{I}}_{bv}}[u(\tau^i_{bv,t})] = V^{\mathrm{IJ}} + \Delta V^{\mathrm{I}}\left(q^{\mathrm{I}}_t + (1-q^{\mathrm{I}}_t)\alpha^{\mathrm{I}}\tau^{\mathrm{I}}_{bv,t}\right) - \frac{1}{2}\left(\tau^i_{bv,t}\right)^2 .$$

Therefore,

$$\begin{aligned}
\Delta U^i = {} & V^{\mathrm{IJ}} + \Delta V^{\mathrm{I}}\left(q^{\mathrm{I}} + (1-q^{\mathrm{I}}_t)\alpha^i \tau^i_{bv}\right) - \frac{1}{2}\left(\tau^i_{bv,t}\right)^2 - V^{\mathrm{IJ}} \\
& - \Delta V^{\mathrm{I}}\left(q^{\mathrm{I}}_t + (1-q^{\mathrm{I}}_t)\left(\alpha^i \tau^i_t\right)\right) + \frac{1}{2}\left(\tau^i_t\right)^2
\end{aligned}$$

$$
\begin{aligned}
&= \Delta V^{\mathrm{I}}(1-q_t^{\mathrm{I}})\alpha^i \tau^i_{bv,t} - \frac{1}{2}\left(\tau^i_{bv,t}\right)^2 - \Delta V^{\mathrm{I}}(1-q_t^{\mathrm{I}})\left(\alpha^i \tau^i_t\right) + \frac{1}{2}\left(\tau^i_t\right)^2 \\
&= \Delta V^{\mathrm{I}}(1-q_t^{\mathrm{I}})\alpha^i (\tau^i_{bv,t} - \tau^i_t) - \frac{1}{2}\left(\tau^{i2}_{bv,t} - \tau^{i2}_T\right) \\
&= \tau^{i2}_{bv} - \tau^i_{bv,t}\tau^i_t - \frac{1}{2}\left(\tau^{i2}_{bv,t} - \tau^{i2}_t\right) \\
&= \frac{1}{2}\tau^{i2}_{bv,t} - \tau^i_{bv,t}\tau^i_t + \frac{1}{2}\tau^{i2}_t \\
&= \frac{1}{2}(\tau^i_{bv,t} - \tau^i_t)^2 . \quad \square
\end{aligned}
$$

Proof of Proposition 3. Recall first that, from equation (6), for each $I \in \mathcal{C}$, $d_t^{\mathrm{I}} = (1-\hat{q}_t^{\mathrm{I}})\omega^{\mathrm{I}}$, and recall also that $b_t^{\mathrm{I}}(q_t^{\mathrm{I}}) = \hat{q}_t^{\mathrm{I}} - q_t^{\mathrm{I}}$ and $b_t^{\mathrm{J}}(q_t^{\mathrm{J}}) = \hat{q}_t^{\mathrm{J}} - (1-q_t^{\mathrm{I}})$. Then

$$
\begin{aligned}
d_t^{\mathrm{I}} - d_t^{\mathrm{J}} &= \left(1 - q_t^{\mathrm{I}} - b_t^{\mathrm{I}}(q_t^{\mathrm{I}})\right)\omega^{\mathrm{I}} - \left(1 - q_t^{\mathrm{J}} - b_t^{\mathrm{J}}(q_t^{\mathrm{I}})\right)\omega^{\mathrm{J}} && \text{(A.18)}\\
&= \left(1 - q_t^{\mathrm{I}} - b_t^{\mathrm{I}}(q_t^{\mathrm{I}})\right)\omega^{\mathrm{I}} - \left(q_t^{\mathrm{I}} - b_t^{\mathrm{J}}(q_t^{\mathrm{I}})\right)\omega^{\mathrm{J}} && \text{(A.19)}\\
&= \left(1 - b_t^{\mathrm{I}}(q_t^{\mathrm{I}})\right)\omega^{\mathrm{I}} + b_t^{\mathrm{J}}\omega^{\mathrm{J}} - q_t^{\mathrm{I}}(\omega^{\mathrm{I}} + \omega^{\mathrm{J}}) && \text{(A.20)}
\end{aligned}
$$

Consider first the stability of the extreme points $q^{\mathrm{I}} = 0$ and $q^{\mathrm{I}} = 1$.

Stability of $\mathbf{q = 0}$. Note that $\frac{\partial \dot{q}_t}{\partial q_t}|_{q_t=0} = d_t^{\mathrm{I}}(q_t^{\mathrm{I}} = 1) - d_t^{\mathrm{J}}(q_t^{\mathrm{I}} = 1)$. Thus $q^{\mathrm{I}} = 0$ stable if and only if $d_t^{\mathrm{I}}(q_t^{\mathrm{I}} = 0) - d_t^{\mathrm{J}}(q_t^{\mathrm{I}} = 0) < 0$, that is

$$
\begin{aligned}
&\left(1 - b_t^{\mathrm{I}}(0)\right)\omega^{\mathrm{I}} + b_t^{\mathrm{J}}(0)\omega^{\mathrm{J}} < 0 \\
&\qquad b_t^{\mathrm{I}}(0) > 1 + b_t^{\mathrm{J}}(0)\frac{\omega^{\mathrm{J}}}{\omega^{\mathrm{I}}}
\end{aligned}
$$

Stability of $\mathbf{q = 1}$. Similarly $q = 1$ is stable if $d_t^{\mathrm{I}}(q_t^{\mathrm{I}} = 1) - d_t^{\mathrm{J}}(q_t^{\mathrm{I}} = 1) > 0$, namely

$$
\begin{aligned}
&-b_t^{\mathrm{I}}(1)\omega^{\mathrm{I}} - \left(1 - b_t^{\mathrm{J}}(1)\right)\omega^{\mathrm{J}} > 0 \\
&\qquad -b_t^{\mathrm{I}}(1)\omega^{\mathrm{I}} < \left(1 - b_t^{\mathrm{J}}(1)\right)\omega^{\mathrm{J}} \\
&\qquad -b_t^{\mathrm{I}}(1) > \left(1 - b_t^{\mathrm{J}}(1)\right)\frac{\omega^{\mathrm{J}}}{\omega^{\mathrm{I}}}
\end{aligned}
$$

We now analyze the possible interior steady states. For a generic $x \in \mathbb{R}$, let $[x]_0^1 := \max\{0, \min\{x, 1\}\}$. Now, $d_t^{\mathrm{I}} - d_t^{\mathrm{J}} = 0$ if and only if $q_t^{\mathrm{I}} = \frac{\omega^{\mathrm{I}}}{\omega^{\mathrm{I}}+\omega^{\mathrm{J}}}\left(1 - b_t^{\mathrm{I}}(q_t^{\mathrm{I}})\right) + \frac{\omega^{\mathrm{J}}}{\omega^{\mathrm{I}}+\omega^{\mathrm{J}}} b_t^{\mathrm{J}}(q_t^{\mathrm{I}})$.

Then, an interior steady state is a $q^{\mathrm{I}} \in (0,1)$ such that

$$
q^{\mathrm{I}} = \left[\frac{\omega^{\mathrm{I}}}{\omega^{\mathrm{I}}+\omega^{\mathrm{J}}}\left(1 - b^{\mathrm{I}}(q^{\mathrm{I}})\right) + \frac{\omega^{\mathrm{J}}}{\omega^{\mathrm{I}}+\omega^{\mathrm{J}}} b^{\mathrm{J}}(q^{\mathrm{I}})\right]_1^0 =: f(q^{\mathrm{I}}).
$$

Note that $f(q^{\mathrm{I}})$ is a self-map and the conditions for the *stability* of $q^{\mathrm{I}} = 0$ and $q^{\mathrm{I}} = 1$ are the conditions for $f(0) = 0$ and $f(1) = 1$ respectively.

Now,

$$
\frac{\partial f(q_t^{\mathrm{I}})}{\partial q_t^{\mathrm{I}}} = \frac{1}{\omega^{\mathrm{I}}+\omega^{\mathrm{J}}}\left(-\omega^{\mathrm{I}}\frac{\partial b_t^{\mathrm{I}}(q_t^{\mathrm{I}})}{\partial q_t^{\mathrm{I}}} + \omega^{\mathrm{J}}\frac{\partial b_t^{\mathrm{J}}(q_t^{\mathrm{I}})}{\partial q_t^{\mathrm{I}}}\right),
$$

and the second derivative:

$$\frac{\partial^2 f(q_t^{\mathrm{I}})}{\partial^2 q_t^{\mathrm{I}}} = \frac{1}{\omega^{\mathrm{I}} + \omega^{\mathrm{J}}}\Big(- \omega^{\mathrm{I}} \frac{\partial^2 b_t^{\mathrm{I}}(q_t^{\mathrm{I}})}{\partial^2 q_t^{\mathrm{I}}} + \omega^{\mathrm{J}} \frac{\partial^2 b^{\mathrm{J}}(q_t^{\mathrm{I}})}{\partial^2 q_t^{\mathrm{I}}}\Big).$$

Note that $\frac{\partial b_t^{\mathrm{I}}(q_t^{\mathrm{I}})}{\partial q_t^{\mathrm{I}}} = \frac{\partial \hat{q}_t^{\mathrm{I}}}{\partial q_t^{\mathrm{I}}} - 1$ and, thus, $\frac{\partial^2 b_t^{\mathrm{I}}(q_t^{\mathrm{I}})}{\partial^2 q_t^{\mathrm{I}}} = \frac{\partial^2 \hat{q}_t^{\mathrm{I}}}{\partial^2 q_t^{\mathrm{I}}}$. Similarly, $\frac{\partial b_t^{\mathrm{J}}(q_t^{\mathrm{I}})}{\partial q_t^{\mathrm{I}}} = \frac{\partial \hat{q}_t^{\mathrm{J}}}{\partial q_t^{\mathrm{I}}} + 1$ and $\frac{\partial^2 b_t^{\mathrm{J}}(q_t^{\mathrm{I}})}{\partial^2 q_t^{\mathrm{I}}} = \frac{\partial^2 \hat{q}_t^{\mathrm{J}}}{\partial^2 q_t^{\mathrm{I}}}$.

Now, if $\frac{\partial^2 \hat{q}_t^{\mathrm{I}}}{\partial^2 q_t^{\mathrm{I}}} > 0$ and $\frac{\partial^2 \hat{q}_t^{\mathrm{J}}}{\partial^2 q_t^{\mathrm{I}}} < 0$, or if $\frac{\partial^2 \hat{q}_t^{\mathrm{I}}}{\partial^2 q_t^{\mathrm{I}}} < 0$ and $\frac{\partial^2 \hat{q}_t^{\mathrm{J}}}{\partial^2 q_t^{\mathrm{I}}} > 0$, then $f(q^{\mathrm{I}})$ does not change concavity since it is always $\frac{\partial^2 f(q_t^{\mathrm{I}})}{\partial^2 q_t^{\mathrm{I}}} < 0$ or always $\frac{\partial^2 f(q_t^{\mathrm{I}})}{\partial^2 q_t^{\mathrm{I}}} > 0$. This, together with $q^{\mathrm{I}} = 0$ and $q^{\mathrm{I}} = 1$ both stable or both unstable, implies uniqueness of the interior steady state. □

Proof of Proposition 4. Consider the optimal socialization effort in (4). Then

$$\begin{aligned}\frac{\partial \tau^i}{\partial q^{\mathrm{I}}} =& \frac{\partial}{\partial q^{\mathrm{I}}}\left[\hat{\alpha}^i (1-\hat{q}^i)\Delta V^{\mathrm{I}}\right] \\ =& \left[-\hat{\alpha}^i \frac{\partial \hat{q}^i}{\partial q^{\mathrm{I}}} + (1-\hat{q}^i)\frac{\partial \hat{\alpha}^i}{\partial q^{\mathrm{I}}}\right]\Delta V^{\mathrm{I}}.\end{aligned}$$

There is cultural substitution if and only if

$$-\hat{\alpha}^i \frac{\partial \hat{q}^i}{\partial q^{\mathrm{I}}} + (1-\hat{q}^i)\frac{\partial \hat{\alpha}^i}{\partial q^{\mathrm{I}}} < 0,$$

$$\frac{\partial \hat{\alpha}^i}{\partial q^{\mathrm{I}}} < \frac{\hat{\alpha}^i}{(1-\hat{q}^i)} \frac{\partial \hat{q}^i}{\partial q^{\mathrm{I}}}. \tag{A.21}$$

Note that, since we have assumed $\frac{\partial \hat{q}^i}{\partial q^{\mathrm{I}}} \geq 0$, then if $\frac{\partial \hat{\alpha}^i}{\partial q^{\mathrm{I}}} < 0$ then cultural substitution is ensured. Note also that condition in (A.21) can be rewrites as

$$\frac{(1-\hat{q}^i)}{\hat{\alpha}^i} \cdot \frac{\partial \hat{\alpha}^i}{\partial q^{\mathrm{I}}} \cdot \frac{\partial q^{\mathrm{I}}}{\partial \hat{q}^i} > 1,$$

so that condition for cultural complementarity immediately follows by reverting the sign of the inequality. □

Proof of Proposition 5. Note first that the equivalence in equation (9) trivially stems from the fact that the updated conjectures on parental efficacy and population shares must be consistent with the Bayesian inference on the transition probabilities (RHS). Let us now focus on the Bayesian updating on the RHS. We have assumed that signals are generated by a probability distribution belonging to the exponential family and that each parent has a conjugated prior. Therefore, to prove the RHS of equation (9), we can exploit results in Diaconis et al. (1979) which proves that the posterior expectation of the mean (i.e., in terms of our model, $\hat{p}_1^{i\mathrm{I}}(\tilde{\tau}_0^i)$) is a linear convex combination of the mean of the prior (i.e., $\hat{p}^{i\mathrm{I}}(\tilde{\tau}_0^i)$) and of the maximum likelihood estimate (i.e., the mean of signals $\bar{s}_{\mathrm{K},0}^i$).

The weighting $\lambda_{\mathrm{K},0}$ is proportional to, and decreasing in the sample size, K. In detail, as proven in Diaconis et al. (1979) (in Remarks at p. 276), $\lambda_{\mathrm{K},0} = \frac{k_0}{k_0+\mathrm{K}}$, where k_0 is a parameter of the conjugate prior and it might be thought of as a prior sample size. Therefore, results for $\mathrm{K} \to \infty$ trivially follows and $\lim_{\mathrm{K}\to\infty} \lambda_{\mathrm{K},0} = 0$. Moreover, for the law of large number $\lim_{\mathrm{K}\to\infty} \bar{s}_{\mathrm{K},0}^i = p^{i\mathrm{I}}(\tilde{\tau}_0)$, that is, the limit of the mean of signals is the mean of the probability distribution, induced by the experimentation effort $\tilde{\tau}_0$, from which signals are generated. Thus, from equation (9) we get $\lim_{\mathrm{K}\to\infty} \hat{p}_1^{i\mathrm{I}}(\tilde{\tau}_0) = p^{i\mathrm{I}}(\tilde{\tau}_0)$ that is

$$\hat{q}_1^i + \hat{\alpha}_1^i \tilde{\tau}_0^i (1-\hat{q}_1^i) = q^{\mathrm{I}} + \alpha^i \tilde{\tau}_0^i (1-q^{\mathrm{I}}).$$

Moreover, if in (9) $\bar{s}^i = p^{i\mathrm{I}}(\tilde{\tau}_0^i)$ we get,

$$\begin{aligned}
\hat{q}_1^i + \hat{\alpha}_1^i \tilde{\tau}_0^i (1-\hat{q}_1^i) =& q^{\mathrm{I}} + \alpha^i \tilde{\tau}_0^{\mathrm{I}} (1-q^{\mathrm{I}}) + \lambda_{\mathrm{K},0}^i \Big(\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^i) \Big), \\
\hat{q}_1^i (1-\hat{\alpha}_1^i \tilde{\tau}_0^i) =& q^{\mathrm{I}} + \alpha^i \tilde{\tau}_0^i (1-q^{\mathrm{I}}) - \hat{\alpha}_1^i \tilde{\tau}_0^i + \lambda_{\mathrm{K},0}^i \Big(\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^i) \Big), \\
\hat{q}_1^i =& \frac{q^{\mathrm{I}} + \alpha^i \tilde{\tau}_0^i (1-q^{\mathrm{I}}) - \hat{\alpha}_1^i \tilde{\tau}_0^i + \lambda_{\mathrm{K},0}^i \Big(\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^i) \Big)}{(1-\hat{\alpha}_1^i \tilde{\tau}_0^i)}.
\end{aligned}$$

Let us subtract q^{I} from both side and we get

$$\begin{aligned}
b_1^i =& \frac{q^{\mathrm{I}} - q^{\mathrm{I}}(1-\hat{\alpha}_1^i \tilde{\tau}_0^i) + \alpha^i \tilde{\tau}_0^i (1-q^{\mathrm{I}}) - \hat{\alpha}_1^i \tilde{\tau}_0^i + \lambda_{\mathrm{K},0}^i \Big(\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^i) \Big)}{(1-\hat{\alpha}_1^i \tilde{\tau}_0^i)} \\
=& \frac{q^{\mathrm{I}} \hat{\alpha}_1^i \tilde{\tau}_0^i + \alpha^i \tilde{\tau}_0^i (1-q^{\mathrm{I}}) - \hat{\alpha}_1^i \tilde{\tau}_0^i + \lambda_{\mathrm{K},0}^i \Big(\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^i) \Big)}{(1-\hat{\alpha}_1^i \tilde{\tau}_0^i)} \\
=& \frac{q^{\mathrm{I}} \tilde{\tau}_0^i (\hat{\alpha}_1^i - \alpha^i) + \tilde{\tau}_0^i \big(\hat{\alpha}_1^i - \alpha^i\big) + \lambda_{\mathrm{K},0}^i \Big(\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^i) \Big)}{(1-\hat{\alpha}_1^i \tilde{\tau}_0^i)} \\
=& \frac{(\alpha^i - \hat{\alpha}_1^{\mathrm{I}}) \tilde{\tau}_0^i (1-q^{\mathrm{I}}) + \lambda_{\mathrm{K},0}^i \Big(\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^{\mathrm{I}}) \Big)}{(1-\hat{\alpha}_1^i \tilde{\tau}_0^i)}.
\end{aligned}$$

Which with an infinite number of signals become

$$b_1^i = \frac{\big(\alpha^i - \hat{\alpha}_1^i\big) \tilde{\tau}_0^i (1-q^{\mathrm{I}})}{(1-\hat{\alpha}_1^i \tilde{\tau}_0^i)}.$$

From equation (9) we get

$$\hat{\alpha}_1^i = \frac{\bar{s}^i + \lambda_{\mathrm{K},0}^i (\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - \bar{s}^i) - \hat{q}_1^i}{\tilde{\tau}_0^i (1-\hat{q}_1^i)}.$$

Since $\bar{s}^i = p^{i\mathrm{I}}(\tilde{\tau}_0^i)$, we can write

$$\hat{\alpha}_1^i = \frac{q^{\mathrm{I}} + \alpha^i \tilde{\tau}_0^i (1-q^{\mathrm{I}}) + \lambda_{\mathrm{K},0}^i (\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - p^{i\mathrm{I}}(\tilde{\tau}_0^i)) - \hat{q}_1^i}{\tilde{\tau}_0^i (1-\hat{q}_1^i)}. \tag{A.22}$$

Thus the socialization effort is

$$\begin{aligned}
\tau^i =& \hat{\alpha}_1^i (1-\hat{q}_1^i) \Delta V^i \\
=& \frac{\bar{s}^i + \lambda_{\mathrm{K},0}^i (\hat{p}_0^{i\mathrm{I}}(\tilde{\tau}_0^i) - \bar{s}^i) - \hat{q}_1^i}{\tilde{\tau}_0^i} \Delta V^i
\end{aligned}$$

$$
\begin{aligned}
&= \tau^i_{bv} - \frac{b^i_1 \Delta V^i}{\tilde{\tau}^i_0} + \frac{\lambda^i_{\mathrm{K},0}(\hat{p}^{i\mathrm{I}}_0(\tilde{\tau}^i_0) - p^{i\mathrm{I}}(\tilde{\tau}^i_0))\Delta V^i}{\tilde{\tau}^i_0} \\
&= \tau^i_{bv} + \left[\lambda^i_{\mathrm{K},0}\Big(\hat{p}^{i\mathrm{I}}_0(\tilde{\tau}^i_0) - p^{i\mathrm{I}}(\tilde{\tau}^i_0)\Big) - b^i_1\right]\frac{\Delta V^i}{\tilde{\tau}^i_0}.
\end{aligned}
$$

Moreover, the conjecture $\hat{q}^i_1$ should be compatible with the locus of point, therefore it means that, from equation (A.22) we have that

$$
\begin{aligned}
\frac{q^{\mathrm{I}} + \alpha^i \tilde{\tau}^i_0(1-q^{\mathrm{I}}) + \lambda^i_{\mathrm{K},0}(\hat{p}^{i\mathrm{I}}_0(\tilde{\tau}^i_0) - p^{i\mathrm{I}}(\tilde{\tau}^i_0)) - \hat{q}^i_1}{\tilde{\tau}^i_0(1-\hat{q}^i_1)} &> 0, \\
q^{\mathrm{I}} + \alpha^i \tilde{\tau}^i_0(1-q^{\mathrm{I}}) + \lambda^i_{\mathrm{K},0}(\hat{p}^{i\mathrm{I}}_0(\tilde{\tau}^i_0) - p^{i\mathrm{I}}(\tilde{\tau}^i_0)) &> \hat{q}^i_1, \\
\hat{q}^i_1 &< \lambda^i_{\mathrm{K},0}\hat{p}^{i\mathrm{I}}_0(\tilde{\tau}^i_0) + (1-\lambda^i_{\mathrm{K},0})p^{i\mathrm{I}}(\tilde{\tau}^i_0). \quad \square
\end{aligned}
$$

Proof of Proposition 6. Given a generic number of experimentation effort we know from equation (9) that the updating of conjectures after one experimentation effort should satisfy:

$$
\hat{p}^{i\mathrm{I}}_n(\tilde{\tau}^i_{n-1}) = \lambda^i_{\mathrm{K},n-1} \cdot \hat{p}^{i\mathrm{I}}_{n-1}(\tilde{\tau}^i_{n-1}) + (1-\lambda^i_{\mathrm{K},n-1}) \cdot \bar{s}^i_n,
$$

given that $\bar{s}^i_n = p^{i\mathrm{I}}(\tilde{\tau}^i_{n-1})$ we can write

$$
\hat{p}^{i\mathrm{I}}_n(\tilde{\tau}^i_{n-1}) = \lambda^i_{\mathrm{K},n-1} \cdot \hat{p}^{i\mathrm{I}}_{n-1}(\tilde{\tau}^i_{n-1}) + (1-\lambda^i_{\mathrm{K},n-1}) \cdot p^{i\mathrm{I}}(\tilde{\tau}^i_{n-1}). \tag{A.23}
$$

If the process converges, at the steady state it holds that

$$
\begin{aligned}
\hat{p}^{i\mathrm{I}}(\tilde{\tau}^i) &= \lambda^i_{\mathrm{K},n-1} \cdot \hat{p}^{i\mathrm{I}}(\tilde{\tau}^i) + (1-\lambda^i_{\mathrm{K},n-1}) \cdot p^{i\mathrm{I}}(\tilde{\tau}^i) \\
\Rightarrow\ \hat{p}^{i\mathrm{I}}(\tilde{\tau}^i) &= p^{i\mathrm{I}}(\tilde{\tau}^i) \\
\Rightarrow\ \hat{q}^i + \hat{\alpha}^i \tau^i (1-\hat{q}^i) &= q^{\mathrm{I}} + \alpha^i \tau^i (1-q^{\mathrm{I}}) \\
\Rightarrow\ \hat{q}^i + \hat{\alpha}^{i2}(1-\hat{q}^i)^2 \Delta V^{\mathrm{I}} &= q^{\mathrm{I}} + \hat{\alpha}^i(1-\hat{q}^i)\Delta V^{\mathrm{I}}\alpha^{\mathrm{I}}(1-q^{\mathrm{I}}) \\
\Rightarrow\ \underbrace{\hat{q}^i - q^{\mathrm{I}}}_{b^i} &= \hat{\alpha}^i(1-\hat{q}^i)\Delta V^{\mathrm{I}}\Big[\alpha^{\mathrm{I}}(1-q^{\mathrm{I}}) - \hat{\alpha}^i(1-\hat{q}^i)\Big].
\end{aligned}
$$

The last equation proves equation (14). Solving it for $\hat{\alpha}^i$ we get the conjectures in equation (15)

$$
\hat{\alpha}^i = \frac{\alpha^{\mathrm{I}}(1-q^{\mathrm{I}})\Delta V^{\mathrm{I}} \pm \sqrt{4(q^{\mathrm{I}} - \hat{q}^i)\Delta V^{\mathrm{I}} + (\alpha^i(1-q^{\mathrm{I}})\Delta V^{\mathrm{I}})^2}}{2(1-\hat{q}^i)\Delta V^{\mathrm{I}}}, \tag{A.24}
$$

then an $\hat{\alpha}^i \in \mathbb{R}_+$ exists if and only if $4(q^{\mathrm{I}} - \hat{q}^i)\Delta V^{\mathrm{I}} + (\tau^i_{bv})^2 \geq 0$. The previous inequality is satisfied if and only if $\hat{q}^i - q^{\mathrm{I}} = b^i \leq \bar{b}^i = \frac{\alpha^{i2}(1-q^{\mathrm{I}})^2\Delta V^{\mathrm{I}}}{4}$. Notice that $\sqrt{4(q^{\mathrm{I}} - \hat{q}^i)\Delta V^{\mathrm{I}} + (\alpha^i(1-q^{\mathrm{I}})\Delta V^{\mathrm{I}})^2} = \sqrt{4\Delta V^{\mathrm{I}}\left(\bar{b}^i - b^i\right)} =: \xi(b^i)$. We can trivially derive equation (16).

To prove the convergence of (A.23) let us write it as

$$
\begin{aligned}
\hat{\alpha}^i_n &= \frac{\lambda^i_{\mathrm{K},n-1} \cdot \hat{p}^{i\mathrm{I}}_{n-1}(\tilde{\tau}^i_{n-1}) + (1-\lambda^i_{\mathrm{K},n-1}) \cdot p^{i\mathrm{I}}(\tilde{\tau}^i_{n-1}) - \hat{q}^i_n}{\tilde{\tau}^i_{n-1}(1-\hat{q}^i_n)} \\
&= \frac{\lambda^i_{\mathrm{K},n-1} \cdot \big(\hat{q}^i_{n-1} + \hat{\alpha}^i_{n-1}\tilde{\tau}^i_{n-1}\big(1-\hat{q}^i_{n-1}\big)\big) + (1-\lambda^i_{\mathrm{K},n-1}) \cdot \big(q^i + \alpha^i \tilde{\tau}^i_{n-1}\big(1-q^i\big)\big) - \hat{q}^i_n}{\tilde{\tau}^i_{n-1}(1-\hat{q}^i_n)}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\lambda_{\mathrm{K},n-1}^{i} \cdot \left(\hat{\alpha}_{n-1}^{i} \left(1 - \hat{q}_{n-1}^{i} \right) \right) + (1 - \lambda_{\mathrm{K},n-1}^{i}) \cdot \left(\alpha^{i} \left(1 - q^{i} \right) \right)}{(1 - \hat{q}_{n}^{i})} \\
&\quad - \frac{b_{n}^{i} - \lambda_{\mathrm{K},n-1}^{i} b_{n-1}^{i}}{\hat{\alpha}_{n-1}^{i} (1 - \hat{q}_{n-1}^{i})(1 - \hat{q}_{n}^{i}) \Delta V^{\mathrm{I}}}.
\end{aligned}
$$

Let us now study the first derivative

$$
\frac{\partial \hat{\alpha}_{n}^{i}}{\partial \hat{\alpha}_{n}^{i}} = \frac{\lambda_{\mathrm{K}} \left(1 - \hat{q}_{n-1}^{i} \right)}{(1 - \hat{q}_{n}^{i})} + \frac{b_{n}^{i} - \lambda_{\mathrm{K}} b_{n-1}^{i}}{\hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i})(1 - \hat{q}_{n}^{i}) \Delta V^{\mathrm{I}}}. \tag{A.25}
$$

If $\lambda_{\mathrm{K},n-1}^{i} = 0$, then (A.25) reads

$$
\frac{\partial \hat{\alpha}_{n}^{i}}{\partial \hat{\alpha}_{n}^{i}} = \frac{b_{n}^{i}}{\hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i})(1 - \hat{q}_{n}^{i}) \Delta V^{\mathrm{I}}},
$$

which is monotone and the sign depends on the sign of the bias. $\frac{\partial \hat{\alpha}_{n}^{i}}{\partial \hat{\alpha}_{n}^{i}} < 1$ if and only if

$$
\begin{aligned}
& b_{n}^{i} < \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i} - \hat{q}_{n}^{i} + \hat{q}_{n-1}^{i} \hat{q}_{n}^{i}) \Delta V^{\mathrm{I}} \\
& \hat{q}_{n}^{i} (1 + \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}}) < q^{\mathrm{I}} + \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}} \\
& \hat{q}_{n}^{i} < \frac{q^{\mathrm{I}} + \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}}}{(1 + \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}})} \\
& \hat{q}_{n}^{i} - q^{\mathrm{I}} < \frac{q^{\mathrm{I}} + \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}}}{(1 + \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}})} - q^{\mathrm{I}} \\
& \hat{q}_{n}^{i} - q^{\mathrm{I}} < \frac{q^{\mathrm{I}} + \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}}}{(1 + \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}})} - \frac{q^{\mathrm{I}} + q^{\mathrm{I}} \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}})}{(1 + \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}})} \\
& \hat{q}_{n}^{i} - q^{\mathrm{I}} < \frac{(1 - q^{\mathrm{I}}) \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}}}{(1 + \hat{\alpha}_{n-1}^{i2} (1 - \hat{q}_{n-1}^{i}) \Delta V^{\mathrm{I}})} = 1 - q^{\mathrm{I}} \\
& \hat{q}_{n}^{i} < 1,
\end{aligned}
$$

always true. Therefore, having infinite signals after each experimentation effort is a sufficient condition for having convergence. □

References

Acemoglu, D., Jackson, M.O., 2014. History, expectations, and leadership in the evolution of social norms. Rev. Econ. Stud. 82 (2), 423–456.

Alesina, A., Miano, A., Stantcheva, S., 2018. Immigration and redistribution. National Bureau of Economic Research. No. w24733.

Bandura, A., 1982. Self-efficacy mechanism in human agency. Am. Psychol. 37 (2), 122–147.

Bandura, A., 1993. Perceived self-efficacy in cognitive development and functioning. Educ. Psychol. 28 (2), 117–148.

Bandura, A., Freeman, W., Lightsey, R., 1999. Self-Efficacy: The Exercise of Control. Springer.

Battigalli, P., 1987. Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali. Unpublished undergraduate dissertation, Bocconi University, Milano.

Battigalli, P., Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., 2015. Self-confirming equilibrium and model uncertainty. Am. Econ. Rev. 105 (2), 646–677.

Battigalli, P., De Vito, N., Catonini, E., 2021. Game Theory: Analysis of Strategic Thinking. Lecture Notes. Bocconi University.
Battigalli, P., Francetich, A., Lanzani, G., Marinacci, M., 2019. Learning and self-confirming long-run biases. J. Econ. Theory 183, 740–785.
Battigalli, P., Gilli, M., Molinari, M.C., 1992. Learning and convergence learning and convergence to equilibrium in repeated strategic interactions: an introductory survey. Res. Econ. 46 (3–4), 335–377.
Battigalli, P., Lanzani, G., 2020. A note on self-confirming equilibrium and stochastic control. Mimeo.
Belsky, J., Jaffee, S., 2006. The multiple determinants of parenting. In: Cicchetti, D., Cohen, D.J. (Eds.), Developmental Psychopathology: Vol. 3: Risk, Disorder and Adaptation. John Wiley & Sons Inc, pp. 38–85, chapter 2.
Belsky, J., Robins, E., Gamble, W., 1984. The determinants of parental competence. In: Lewis, M. (Ed.), Beyond The Dyad. In: Genesis of Behavior, vol. 4. Springer, Boston, MA, pp. 251–279.
Bisin, A., Topa, G., Verdier, T., 2004. Cooperation as a transmitted cultural trait. Ration. Soc. 16 (4), 477–507.
Bisin, A., Verdier, T., 2001. The economics of cultural transmission and the dynamics of preferences. J. Econ. Theory 97 (2), 298–319.
Bisin, A., Verdier, T., 2011. The economics of cultural transmission and socialization. In: Benhabib, Jess, Bisin, Alberto, Jackson, Matt (Eds.), Handbook of Social Economics, vol. 1. Elsevier, chapter 9.
Bisin, A., Verdier, T., 2017. On the joint evolution of culture and institutions. Technical report. National Bureau of Economic Research.
Büchel, B., Hellmann, T., Pichler, M.M., 2014. The dynamics of continuous cultural traits in social networks. J. Econ. Theory 154, 274–309.
Cavalli-Sforza, L.L., Feldman, M.W., 1981. Cultural Transmission and Evolution: A Quantitative Approach. Princeton University Press.
Cheung, M.-W., Wu, J., 2018. On the probabilistic transmission of continuous cultural traits. J. Econ. Theory 174, 300–323.
Coleman, P.K., Karraker, K.H., 2000. Parenting self-efficacy among mothers of school-age children: conceptualization, measurement, and correlates. Fam. Relat. 49 (1), 13–24.
de Montigny, F., Lacharité, C., 2005. Perceived parental efficacy: concept analysis. J. Adv. Nurs. 49 (4), 387–396.
Della Lena, S., Dindo, P., 2019. On the evolution of norms and strategic environments. In: University Ca'Foscari of Venice, Dept. of Economics Research Paper Series, vol. 16.
Diaconis, P., Ylvisaker, D., et al., 1979. Conjugate priors for exponential families. Ann. Stat. 7 (2), 269–281.
Doepke, M., Zilibotti, F., 2017. Parenting with style: altruism and paternalism in intergenerational preference transmission. Econometrica 85 (5), 1331–1371.
Doepke, M., Zilibotti, F., 2019. Love, Money, and Parenting: How Economics Explains the Way We Raise Our Kids. Princeton University Press.
Donovan, W.L., Leavitt, L.A., 1985. Simulating conditions of learned helplessness: the effects of interventions and attributions. Child Dev. 56 (3), 594–603.
Donovan, W.L., Leavitt, L.A., Walsh, R.O., 1997. Cognitive set and coping strategy affect mothers' sensitivity to infant cries: a signal detection approach. Child Dev. 68 (5), 760–772.
Easley, D., Kiefer, N.M., 1988. Controlling a stochastic process with unknown parameters. Econometrica 85 (5), 1045–1064.
Fudenberg, D., Kreps, D.M., 1995. Learning in extensive-form games I. self-confirming equilibria. Games Econ. Behav. 8 (1), 20–55.
Fudenberg, D., Levine, D.K., 1993. Self-confirming equilibrium. Econometrica 61 (3), 523–545.
Gilli, M., 1999. Adaptive learning in imperfect monitoring games. Rev. Econ. Dyn. 2 (2), 472–485.
Mash, E.J., Johnston, C., 1983. Parental perceptions of child behavior problems, parenting self-esteem, and mothers' reported stress in younger and older hyperactive and normal children. J. Consult. Clin. Psychol. 51 (1), 86–99.
Nteta, T.M., Wallsten, K.J., 2012. Preaching to the choir? Religious leaders and American opinion on immigration reform. Soc. Sci. Q. 93 (4), 891–910.
Panebianco, F., 2014. Socialization networks and the transmission of interethnic attitudes. J. Econ. Theory 150, 583–610.
Prummer, A., Siedlarek, J.-P., 2017. Community leaders and the preservation of cultural traits. J. Econ. Theory 168, 143–176.
Spiro, D., 2020. Multigenerational transmission of culture. J. Econ. Theory 188, 105037. https://doi.org/10.1016/j.jet.2020.105037.
Tabellini, G., 2008. The scope of cooperation: values and incentives. Q. J. Econ. 123 (3), 905–950.
Unger, D.G., Wandersman, L.P., 1985. Social support and adolescent mothers: action research contributions to theory and application. J. Soc. Issues 41 (1), 29–45.

Van Bakel, H.J., Riksen-Walraven, J.M., 2002. Parenting and development of one-year-olds: links with parental, contextual, and child characteristics. Child Dev. 73 (1), 256–273.

Vaughan, D., 2013. Nurture vs. nurture: Endogenous parental and peer effects and the transmission of culture. Technical report, Working Papers. Banco de Mexico (No. 2013-04).

Verdier, T., Zenou, Y., 2015. The role of cultural leaders in the transmission of preferences. Econ. Lett. 136, 158–161.

Verdier, T., Zenou, Y., 2018. Cultural leader and the dynamics of assimilation. J. Econ. Theory 175, 374–414.

Wells-Parker, E., Miller, D.I., Topping, J.S., 1990. Development of control-of-outcome scales and self-efficacy scales for women in four life roles. J. Pers. Assess. 54 (3–4), 564–575.