

On the Probabilistic Transmission of Continuous Cultural Traits*

Man-Wah Cheung[†]

School of Economics

Shanghai University of Finance and Economics

Jiabin Wu

Department of Economics

University of Oregon

February 6, 2017

Abstract

This paper proposes a framework that generalizes the discrete cultural transmission model of Bisin and Verdier (2001) to a continuous trait setting. We define the cultural distance between two agents as the distance of their traits in the trait space, and model an agent's cultural intolerance towards another agent as an increasing function of their cultural distance. This captures people's general tendencies of evaluating culturally more distant people with stronger biases. The resulting cultural evolutionary dynamic can be viewed as a continuous imitative dynamic (as studied in Cheung (2016)) in a population game in which a player's payoff is equal to the aggregate cultural intolerance he has towards other agents. We use cultural intolerance to define *cultural substitutability* in the continuous trait setting. We find that as in Bisin and Verdier (2001), cultural substitutability is the key to cultural heterogeneity. Furthermore, the curvature of the cultural intolerance function plays an important role in determining the long-run cultural phenomena. In particular, when the cultural intolerance function is convex, only the most extremely polarized state is a stable limit point.

Keywords: Cultural transmission, Continuous trait space, Cultural evolution, Imitative Dynamic, Polarization.

JEL Classifications: A14, C72, C73, D10, Z13

*We thank Bill Sandholm, Joseph McMurray, Farid Toubal for helpful discussions and suggestions. Please direct correspondence to either author. Cheung: jennymwcheung@gmail.com, Wu: jwu5@uoregon.edu.

[†]Financial support from National Natural Science Foundation of China (No. 71601102) is gratefully acknowledged.

1 Introduction

During the past decade, inter-generational cultural transmission has become an important subject in economics (see Bisin and Verdier (2010) for an extensive survey). Economists realize that to have a better understanding of economic behavior and outcomes, it is essential to understand how economically relevant cultural traits such as preferences, attitudes, opinions and beliefs are formed and transmitted across generations. Bisin and Verdier (2001) propose a seminal utility-maximizing framework for studying inter-generational transmission of binary cultural traits, which is later generalized by Bisin et al (2009) and Montgomery (2010) to multiple-trait settings. While Bisin and Verdier (2001)'s model is widely used to study socioeconomic phenomena such as corruption (Hauk and Saez-Marti (2002)), religious intermarriage (Bisin et al (2004)) and cooperation (Tabellini (2008)), the discrete nature of the model may limit its applicability to the study of traits such as risk preferences, patience and interethnic attitudes (Panebianco (2014) and Buechel et al (2014)), which are naturally modeled as continuous.

In this paper, we study a cultural transmission model that generalizes Bisin and Verdier (2001)'s utility-maximizing framework to a continuous trait setting. The cultural transmission process we consider is of probabilistic nature as in Bisin and Verdier (2001). Parents can costly exert effort to increase the probability of successful inculcation of their own traits to their children. If inculcation fails, a child adopts the trait from a randomly chosen role model in the society. We also maintain the important assumption called “imperfect empathy” in Bisin and Verdier (2001). That is, parents care about their children's adopted cultural traits but they are biased in that they are intolerant towards other traits (*cultural intolerance*).¹

We find that the resulting cultural evolutionary dynamic can be viewed as a continuous imitative dynamic (cf. Cheung (2016))² of a population game in which a player's payoff is equal to the aggregate cultural intolerance he has towards other agents. Montgomery (2010) shows that in the finite trait setting, the resulting cultural evolutionary dynamic can be viewed as the discrete replicator dynamic if the probability of successful inculcation is a linear function of effort and the effort cost function is quadratic. Our result thus generalizes Montgomery (2010)'s result from the finite trait setting to the continuous trait setting, and from the replicator dynamic to the more general imitative dynamics. Using the properties of imitative dynamics, we show that cultural substitutability leads to the commonly observed phenomenon of cultural heterogeneity as in Bisin and Verdier (2001).

¹Formally, imperfect empathy means that parents evaluate their children's payoffs through their own (the parental) utility functions. This indicates that parents desire their children to adopt their (the parents') traits.

²Note that the replicator dynamic is a special case of imitative dynamics.

Our model contrasts to several recent works including Bisin and Topa (2003), Doepke and Zilibotti (2008), Pichler (2010), Vaughan (2010), Buechel et al (2014), and Panebianco (2014), who consider continuous cultural trait transmission models based on Cavalli-Sforza and Feldman (1973)'s non-probabilistic framework: a child's trait is a weighted average of his parent's trait and the mean value of the trait in the society. A number of authors argue that (e.g., Buechel et al, 2014, Remark 1) the long run cultural heterogeneity result of Bisin and Verdier (2001) disappears when one takes a continuous trait framework like the linear weighting one of Cavalli-Sforza and Feldman (1973) or Boyd and Richerson (1985). Indeed in such models, vertical and oblique transmissions interact in such a way that one obtains a standard mean reverting linear process, leading naturally to cultural homogeneity in the long run. Instead, our result shows that when one allows for probabilistic transmission based on some endogenous strategic paternalistic interaction between vertical and oblique transmissions (as in the initial spirit of Bisin and Verdier (2001)), then cultural substitutability is sufficient to generate cultural heterogeneity in the long run, even when the trait space is continuous.

Next, we equip the model with a novel feature: cultural intolerance as an increasing function of *cultural distance*, i.e., the distance between two traits in the trait space.³ This captures people's general tendencies of evaluating culturally more distant people with stronger biases, which is consistent with the recent findings in the literature of empirical cultural economics (e.g., Guiso, Sapienza and Zingales (2006, 2009), Spolaore and Waacziarg (2009, 2013)) that people are less comfortable with culturally more distant people.⁴

We use cultural intolerance to define cultural substitutability in the continuous trait setting. With our definition of cultural substitutability, Bisin and Verdier (2001)'s long run heterogeneity result extends to the continuous trait setting. Furthermore, we find that the curvature of the cultural intolerance function is crucial for determining the prediction of the cultural evolutionary dynamic. In particular, when the cultural intolerance function is convex, only those states that consist of no more than two traits are limit points of the cultural evolutionary dynamic. And among those states, only the most extremely polarized one is Lyapunov stable. The intuition is as follows. Parents make inculcation efforts that increase with aggregate cultural intolerance. Thus traits experiencing the highest aggregate cultural intolerance grow most quickly, which alters the

³We define the continuous trait space to be the unit interval.

⁴As argued in Spolaore and Wacziarg (2009, 2013), differences in cultural traits can hinder development through a barrier effect. Less closely related societies are less likely to learn from each other, and so the diffusion of complex technological and institutional innovations is limited. Guiso, Sapienza and Zingales (2009) find that in the context of economic exchange, people from culturally more similar countries tend to trust each other more and breach contracts less frequently. See also Felbermayr and Toubal (2010) and Hahn (2014).

profile of aggregate cultural intolerance. Under convexity of the cultural intolerance function, the highest aggregate cultural intolerance is experienced by the two most extreme traits. As a result, the dynamic always has the tendency to move away from the current state as long as there are positive masses on both of the most extreme traits, and the population eventually splits into two opposing groups with these two most extreme traits. This shows that convexity of the cultural intolerance function is a sufficient condition for polarization.

When the cultural intolerance function is strictly concave, polarization is no longer Lyapunov stable under the cultural evolutionary dynamic. Moreover, the set of limit points under a strictly concave cultural intolerance function is a strict superset of the set of limit points under a convex cultural intolerance function. That is, the cultural evolutionary dynamic under a strictly concave cultural intolerance function can lead to more diversified population states. Intuitively, since the increment in cultural intolerance decreases as cultural distance increases, one can always find population states in which some traits different from the two most extreme traits are associated with aggregate cultural intolerances higher than those of the two most extreme traits. This result implies that non-concavity of the cultural intolerance function serves as a necessary condition for polarization.

In sum, our model provides a cultural transmission explanation of why a society may or may not be divided into two opposing groups defined by conflicting preferences or beliefs.⁵

The rest of this paper is organized as follows. Section 2 describes the inter-generational cultural transmission mechanism for continuous trait space, and shows that under cultural substitutability, the cultural evolutionary dynamic is equivalent to an imitative dynamic in a population game in which a player's payoff is equal to the aggregate cultural intolerance he has towards other agents. Section 3 shows that as in Bisin and Verdier (2001), cultural substitutability leads to cultural heterogeneity. Section 4 investigates the effects of the curvature of the cultural intolerance function on the long-run behavior of the cultural evolutionary dynamic. Section 5 concludes. Some proofs and definitions omitted from the text are presented in the Appendix.

⁵Polarization is an important empirical phenomenon. For example, "culture wars" in the United States between the liberal left and the conservative right have been a heated political topic for decades (see Hunter (1992)). Polarization also has significant impacts on economics and politics. As argued by Sartori (1966, 1976), public ideological movements towards extremism serves as the basis for the creation of highly polarized multiparty systems. Polarization in people's political beliefs, also known as political polarization, has been investigated in the literature of political economics. For example, Dixit and Weibull (2007) propose an elegant model to illustrate that political polarization can arise from heterogeneous beliefs about the nature of the world even people's preferences are homogeneous. See also McMurray (2015). In addition, confirmation bias in psychology (see Nickerson (1998), Rabin and Schrag (1999)) can also give rise to polarization in beliefs. Yet to our limited knowledge, no work has been done to explain polarization in preferences or other persistent traits.

2 Inter-generational Cultural Transmission

Inter-generational cultural transmission mechanisms are first formally introduced by Cavalli-Sforza and Feldman (1981) and Boyd and Richerson (1985). Building on those works, Bisin and Verdier (2001) introduce a model in which the probability that a child adopts a certain trait is endogenously determined by the parent’s effort. One crucial assumption made by Bisin and Verdier (2001) is that the parents have “imperfect empathy”. That is, a parent’s effort choice decision is motivated by his evaluation of his child’s payoff, and the evaluation is based on the parent’s own utility function (*paternalistic altruism*). Hence, a parent is always biased against other traits (*cultural intolerance*) and wants to inculcate his own trait into his child. It turns out to be optimal for parents with stronger cultural intolerance exert higher efforts in inculcation. Nevertheless, the parent to child transmission is probabilistic, so there is still some chance that the child adopts a trait that is different from his parent’s. In this section, we follow the same spirit of Bisin and Verdier (2001) and build a cultural transmission mechanism, but on continuous trait space.

2.1 The model

Consider a population of agents of unit mass. Each agent in the population has a trait from set $T = [0, 1]$. If the trait space T were finite, then the population state at a particular time could be described by a vector in $\mathbb{R}^{|T|}$,⁶ as in Bisin and Verdier (2001), Bisin et al (2009), and Montgomery (2010). On the other hand, for the case of continuous trait space we consider, the population state is a distribution of traits over T and thus is described by a probability measure over T .

At each time t , agents in the population are selected uniformly at random, and the selected agent becomes a parent and bears a child. Assume that the child does not have a defined trait, and the parent has to decide how much effort to exert to inculcate his own trait into his child. Such socialization within the family is called “direct vertical” socialization. If the parent fails to inculcate his own trait in his child, the child will randomly search for a role model in the society and adopt the trait of the role model. Such socialization by the society is called “oblique” socialization.⁷ After the child’s trait is formed, the parent is replaced by the child in the current population.

Consider the decision of a parent with trait $z \in T$ about how much effort to put into inculcation. Let V_{zy} denote the parent’s assessment of his child’s payoff if his child finally adopts trait $y \in T$. Assume that $V_{zz} \geq V_{zy}$ for any $y \in T$. This reflects the “imperfect empathy” assumption in Bisin and Verdier (2001): a parent always believes that his child would be better off to adopt his own

⁶ $|A|$ denotes the number of elements in set A .

⁷Refer to Cavalli-Sforza and Feldman (1981), or Bisin and Verdier (2001) for the terminology.

trait. Without loss of generality, we assume that $V_{zy} \in [0, 1]$ for all $z, y \in T$.

Denote by $\Delta_{zy} := V_{zz} - V_{zy}$ the *cultural intolerance* a z -parent has towards trait $y \in T$. We assume that V_{zy} is continuous in z and y , and hence Δ_{zy} is continuous in z and y . This assumption captures the following two important features of our cultural transmission model with continuous trait space, which are absent when the trait space is finite:

- i) The continuity of V_{zy} in y means that the evaluation of utility for the other trait y by a z -parent is continuous in y .
- ii) The continuity of V_{zy} in z means that the intolerance structure/pattern of a z -parent is continuous in the trait z of the parent.

One can immediately see that $\Delta_{zz} = 0$ for any $z \in T$, which means that the cultural intolerance a z -parent has towards his own trait is always zero. Also, we have $\Delta_{zy} \in [0, 1]$ for any $z, y \in T$.

Let \mathcal{B} be the Borel σ -algebra on T . Denote by $\mathcal{M}_1^+(T)$ the space of probability measures on (T, \mathcal{B}) . A *population state* is a distribution of traits over T and is described by a probability measure $\mu \in \mathcal{M}_1^+(T)$. The *aggregate cultural intolerance* of a z -parent at population state μ is

$$\Delta_z(\mu) = \int_{y \in T} \Delta_{zy} \mu(dy).$$

Denote by $e \in \mathbb{R}_+$ the inculcation effort exerted by a parent. Let $d : \mathbb{R}_+ \rightarrow [0, 1]$ be the probability of success function, where $d(e)$ is the probability that a parent successfully inculcates his own trait in his child when he exerts effort e .

As in Bisin and Verdier (2001), we allow the inculcation effort depend on the distribution of traits. Let $e_z(\mu)$ be the effort exerted by a z -parent at population state $\mu \in \mathcal{M}_1^+(T)$. Then the probability of successful inculcation by a z -parent at state μ is $d(e_z(\mu))$.

The cultural evolutionary dynamic is characterized by the following differential equation on $\mathcal{M}_1^+(T)$:

$$\dot{\mu}(A) = \int_{y \in A} \int_{z \in T \setminus A} (1 - d(e_z(\mu))) \mu(dz) \mu(dy) - \int_{y \in A} \int_{z \in T \setminus A} (1 - d(e_y(\mu))) \mu(dz) \mu(dy), \quad (\text{CED})$$

for all $A \in \mathcal{B}$. The term $\dot{\mu}(A)$ is the rate of change in the mass of agents with traits in set A . The first term on the RHS of (CED) is the “inflow” of children whose parents’ traits are not in set A but who themselves adopt traits in set A , and the second term is the “outflow” of children whose parents’ traits are in set A but who themselves adopt traits not in set A .

2.2 Cultural Substitutability

We use cultural intolerance to define cultural substitutability as follows.

Definition 1 *Direct vertical cultural transmission and oblique cultural transmission are cultural substitutes for z -parent if*

$$d(e_z(\mu)) = D(\Delta_z(\mu)),$$

where D is a Lipschitz continuous and strictly increasing function. That is, the probability of successful inculcation by a z -parent depends only on the aggregate cultural intolerance of the parent, and increases with the aggregate cultural intolerance.

Intuitively, we say that direct vertical cultural transmission acts as a cultural substitute to oblique cultural transmission whenever parents socialize their children more intensely the higher aggregate cultural intolerance they experience.

Suppose that direct vertical cultural transmission and oblique cultural transmission are cultural substitutes for agents of any trait. Denote by $\mathcal{M}(T)$ the space of finite signed measures on (T, \mathcal{B}) . Let $V : \mathcal{M}_1^+(T) \rightarrow \mathcal{M}(T)$ be defined by the RHS of the cultural evolutionary dynamic (CED), that is, $V(\mu) \in \mathcal{M}(T)$ is defined by

$$V(\mu)(A) := \int_{y \in A} \int_{z \in T \setminus A} (1 - d(e_z(\mu))) \mu(dz) \mu(dy) - \int_{y \in A} \int_{z \in T \setminus A} (1 - d(e_y(\mu))) \mu(dz) \mu(dy),$$

for $\mu \in \mathcal{M}_1^+(T)$ and $A \in \mathcal{B}$. Then for any $A \in \mathcal{B}$,

$$\begin{aligned} V(\mu)(A) &= \int_{y \in A} \int_{z \in T} (d(e_y(\mu)) - d(e_z(\mu))) \mu(dz) \mu(dy) \\ &= \int_{y \in A} \int_{z \in T} (D(\Delta_y(\mu)) - D(\Delta_z(\mu))) \mu(dz) \mu(dy) \\ &= \int_{y \in A} D(\Delta_y(\mu)) \mu(dy) - \mu(A) \int_{z \in T} D(\Delta_z(\mu)) \mu(dz). \end{aligned} \tag{1}$$

To further understand the cultural evolutionary dynamic, consider a random matching environment in which a unit mass of agents are randomly matched in pairs to play a two-player symmetric game with strategy set T and payoff function Δ_{xy} . That is, Δ_{xy} is the single match payoff of an agent playing strategy x against an opponent playing strategy y . Define

$$F_x(\mu) := \int_{y \in T} \Delta_{xy} \mu(dy), \quad \text{for } \mu \in \mathcal{M}_1^+(T) \text{ and } x \in T. \tag{2}$$

Then $F_x(\mu)$ is the expected payoff of an agent playing pure strategy $x \in T$ at population state $\mu \in \mathcal{M}_1^+(T)$. Note that $F(\mu)$ is the payoff profile over T at state μ , and we may just denote it as $F(\mu)$. From Example 1 of Cheung (2014), F , as a map from $\mathcal{M}_1^+(T)$ to $\mathcal{C}_b(T)$, where $\mathcal{C}_b(T)$ is the space of bounded continuous functions on T with the supremum norm, defines a population game.⁸

From the definition of F in (2), $F_z(\mu)$ is simply the aggregate cultural intolerance of a z -parent at population state μ , i.e., $F_z(\mu) = \Delta_z(\mu)$. Together with (1), we have for any $A \in \mathcal{B}$,

$$V(\mu)(A) = \int_{y \in A} D(F_y(\mu)) \mu(dy) - \mu(A) \int_{z \in T} D(F_z(\mu)) \mu(dz). \quad (3)$$

Next, we will show that the cultural evolutionary dynamic under cultural substitutability is an imitative dynamic for the population game F defined by (2). Cheung (2016) shows that monotonicity conditions for imitative dynamics are most naturally expressed in terms of Radon-Nikodym derivatives. From (3), we have for any $A \in \mathcal{B}$,

$$\mu(A) = 0 \quad \Rightarrow \quad V(\mu)(A) = 0.$$

So for any $\mu \in \mathcal{M}_1^+(T)$, $V(\mu) \ll \mu$, and hence the Radon-Nikodym derivative $\frac{dV(\mu)}{d\mu}$ exists.⁹ In particular,

$$\frac{dV(\mu)}{d\mu}(y) = D(F_y(\mu)) - \int_{z \in T} D(F_z(\mu)) \mu(dz). \quad (4)$$

Since $D(\cdot)$ is a strictly increasing function and the second term on the RHS of (4) is independent of y , we have for any $x, y \in T$,

$$\frac{dV(\mu)}{d\mu}(y) > \frac{dV(\mu)}{d\mu}(x) \quad \Longleftrightarrow \quad F_y(\mu) > F_x(\mu).$$

Therefore, the cultural evolutionary dynamic under cultural substitutability satisfies the payoff monotonicity condition (PM) in Cheung (2016), and hence is equivalent to an imitative dynamic for the population game F defined by (2).¹⁰

⁸A population game is identified with a map from $\mathcal{M}_1^+(T)$ to $\mathcal{C}_b(T)$ that is continuous with respect to the weak topology (cf. Cheung, 2014, Section 2.1). A sequence of measures $\mu_n \in \mathcal{M}_1^+(T)$ converges weakly to $\mu \in \mathcal{M}_1^+(T)$, written $\mu_n \xrightarrow{w} \mu$, if $\int_S f d\mu_n \rightarrow \int_S f d\mu$ for all $f \in \mathcal{C}_b(T)$. A map $F : \mathcal{M}_1^+(T) \rightarrow \mathcal{C}_b(T)$ is continuous with respect to the weak topology if $F(\mu_n) \rightarrow F(\mu)$ (in the supremum norm) for any sequence $\{\mu_n\} \subseteq \mathcal{M}_1^+(T)$ such that $\mu_n \xrightarrow{w} \mu$.

⁹See, e.g., Rudin (1987) for the relationship between absolute continuity of measures and Radon-Nikodym derivatives.

¹⁰Cf. Cheung (2016) for the definition of imitative dynamics for games with continuous strategy space.

From Theorem 2 of Cheung (2016) and the discussion after that theorem, the cultural evolutionary dynamic under cultural substitutability is well-defined if $D(\cdot)$ is Lipschitz continuous. We summarize the result in the following theorem.

Theorem 1 (The Cultural Evolutionary Dynamic under Cultural Substitutability is Well-defined)¹¹ *Under cultural substitutability, the cultural evolutionary dynamic is equivalent to an imitative dynamic for the population game F defined by (2), and is well-defined (i.e., for any initial distribution of traits, solutions for the dynamic exist and are unique). Furthermore, solutions to the dynamic are continuous in the initial distribution of traits.*

Thus as in Montgomery (2010), we can reinterpret our cultural evolutionary dynamic as an imitative dynamic for the population game F defined by (2), and can use the game-theoretic approach to study the dynamic. Under the reinterpretation, “traits” become “pure strategies” (elements of T), the “cultural intolerance that a z -parent has towards trait y ” becomes the “single match payoff of an agent playing pure strategy z against an opponent playing pure strategy y ” (Δ_{zy}), and the “aggregate cultural intolerance of a z -parent” becomes the “expected payoff of an agent playing pure strategy z in the random matching game” ($\int_{y \in T} \Delta_{zy} \mu(dy)$).

2.3 Endogenous Cultural Transmission Mechanism

Following Bisin and Verdier (2001), we now lay out a cultural transmission mechanism in which the choices of inculcation effort are endogenously determined.

Let $c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be the effort cost function, i.e., $c(e)$ is the cost associated with effort e . We maintain the following assumptions on the probability of success function $d(\cdot)$ and the effort cost function $c(\cdot)$:

- i) $d : \mathbb{R}_+ \rightarrow [0, 1]$ and $c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ are twice-differentiable;
- ii) $d(0) = 0$, $d' > 0$ and $d'' \leq 0$;
- iii) $c(0) = 0$, $c'(0) = 0$, $c' \geq 0$ and $c'' > 0$.

At state $\mu \in \mathcal{M}_1^+(T)$, a z -parent solves the following maximization problem:

$$\max_{e_z} d(e_z)V_{zz} + (1 - d(e_z)) \int_{y \in T} V_{zy} \mu(dy) - c(e_z). \quad (\text{M}_z)$$

¹¹Cf. Cheung, 2016, Theorem 2. By putting $r_{zy}(\pi) = 1 - D(\pi(z))$ as the conditional switch rate from z to y under payoff profile π into Theorem 2 of Cheung (2016), we obtain the result.

The z -parent evaluates the payoff of his child from his own point of view (i.e., based on his own utility function), taking into account both the possibility that he successfully inculcates his own trait into his child, as well as the possibility that he fails to do so and his child adopts the trait of a randomly selected role model in the population. We have the following proposition.

Proposition 1 *At state $\mu \in \mathcal{M}_1^+(T)$, the optimal effort exerted by a z -parent in maximization problem (M_z) is*

$$e_z^*(\mu) = \left(\frac{c'}{d'}\right)^{-1}(\Delta_z(\mu)) \geq 0,$$

which is bounded. In particular, the z -parent's optimal effort $e_z^(\mu)$ strictly increases with his aggregate cultural intolerance $\Delta_z(\mu)$.*

Proof. See Appendix A.2.

Intuitively, the greater a z -parent's aggregate cultural intolerance is, the better off the z -parent believes his child would be if his child adopts trait z rather than a randomly chosen one, and hence the more effort the z -parent exerts.

By Proposition 1,

$$d(e_z^*(\mu)) = d\left(\left(\frac{c'}{d'}\right)^{-1}(\Delta_z(\mu))\right).$$

Since $d\left(\left(\frac{c'}{d'}\right)^{-1}(\cdot)\right)$ is a Lipschitz continuous and strictly increasing function, cultural substitutability is satisfied. That is, the endogenous cultural transmission mechanism leads to cultural substitutability. Thus by Theorem 1, the cultural evolutionary dynamic under the endogenous cultural transmission mechanism is equivalent to an imitative dynamic for the population game F defined by (2).

3 Cultural Heterogeneity

3.1 Stability and Convergence for Imitative Dynamics

To study stability under the cultural evolutionary dynamic, we need to consider “closeness” and “neighborhoods” of population states, which depend on the choice of topology for the space of measures. We find it most appropriate to use the weak topology on $\mathcal{M}_1^+(T)$ to study dynamic

stability, as we now explain.¹²

Suppose that two traits $x, y \in [0, 1]$ are very close to each other in the trait space. If we adopt the strong topology, the monomorphic state corresponding to x and that corresponding to y are considered to have the maximum distance apart in the state space. But in the weak topology, the two monomorphic states described above are regarded to be close to each other.¹³ Therefore, the weak topology appears to provide a more natural definition of “closeness” of population states.

Some terms standard in dynamical systems theory and in evolutionary game theory (e.g., Lyapunov stability, ω -limit points) are used below. Their formal definitions can be found in Appendix A.1.

Let $F : \mathcal{M}_1^+(T) \rightarrow \mathcal{C}_b(T)$ be a population game. A population state $\mu^* \in \mathcal{M}_1^+(T)$ is a *Nash equilibrium* (NE) of F if

$$F_y(\mu^*) \leq F_z(\mu^*), \quad \forall z \in \text{supp}(\mu^*), \quad \forall y \in T. \quad (\text{NE})$$

We denote by $NE(F)$ the set of Nash equilibria of F , which can be described as

$$NE(F) := \left\{ \mu \in \mathcal{M}_1^+(T) : x \in \text{supp}(\mu) \Rightarrow F_x(\mu) = \max_{y \in T} F_y(\mu) \right\}.$$

Define the set of *restricted equilibria*¹⁴ of F by

$$RE(F) := \left\{ \mu \in \mathcal{M}_1^+(T) : x \in \text{supp}(\mu) \Rightarrow F_x(\mu) = \max_{y \in \text{supp}(\mu)} F_y(\mu) \right\}.$$

In words, μ is a restricted equilibrium of F if it is a Nash equilibrium of a restricted version of F in which only strategies in the support of μ can be played. In particular, all Nash equilibria are restricted equilibria. An imitative dynamic in F has the following properties:¹⁵

P1) the set of rest points of the dynamic coincides with $RE(F)$;

P2) any non-Nash rest points are not Lyapunov stable under the dynamic.

¹²See also Oechssler and Riedel (2002) for an extensive discussion on the issues of choosing between the strong topology and the weak topology.

¹³See Appendix A.1.

¹⁴Cf. Sandholm, 2010, Section 5.4.6.

¹⁵Cf. Cheung, 2016, Propositions 2 and 3.

3.2 Cultural Heterogeneity

One powerful conclusion drawn from Bisin and Verdier (2001) is that cultural substitutability is sufficient to generate cultural heterogeneity, a phenomenon commonly observed in the real world. Recently, Panebianco (2014) and Buechel et al (2014) claim that cultural substitutability alone can no longer lead to cultural heterogeneity when the trait space is continuous. However, the models they consider are based on the non-probabilistic cultural transmission model of Cavalli-Sforza and Feldman (1973). In this section, we show that cultural heterogeneity is preserved in probabilistic cultural transmission model on continuous trait space.

Proposition 2 *Any monomorphic state is not a Nash equilibrium of the population game F defined by (2), and thus is not Lyapunov stable under the cultural evolutionary dynamic.*

Proof. See Appendix A.3.

Intuitively, Proposition 2 shows that cultural homogeneity is unstable under the cultural evolutionary dynamic. The reason is as follows. Consider the induced population game F defined by (2). In a monomorphic state, every agent is using the same strategy, which gives them zero payoff. Meanwhile, using any other strategy gives them a positive payoff. Hence, agents have incentive to deviate. This implies that any monomorphic state cannot be a Nash equilibrium.

Proposition 2 shows that the long run cultural heterogeneity result of Bisin and Verdier (2001) extends to the continuous trait setting: cultural substitutability leads to cultural heterogeneity.

4 Cultural Polarization

4.1 Symmetric Cultural Intolerance and Potential Games

In the remainder of the paper, we focus on cultural intolerance structures that satisfy $\Delta_{xy} = \Delta_{yx}$ for any $x, y \in T$. This means that parents with different traits have symmetric cultural intolerance towards each other. Such intolerance structures resemble the payoff structures for potential games (in particular, doubly symmetric games) with continuous strategy space (cf. Oechssler and Riedel (2002), Cheung (2014, 2016)).¹⁶ When $\Delta_{xy} = \Delta_{yx}$ for any $x, y \in T$, the population game F defined by (2) is a potential game.¹⁷

¹⁶See Montgomery (2010) for a discussion on cultural intolerance structures that resemble the payoff structures for potential games with finite number of strategies when the trait space is finite.

¹⁷See Cheung (2014, 2016) for the formal definition of potential games in continuous strategy settings.

Below are some useful stability and convergence results on imitative dynamics in potential games. Consider an imitative dynamic in potential game F . Suppose that the initial condition is $\xi \in \mathcal{M}_1^+(T)$.¹⁸ The ω -limit set $\omega(\xi)$ is the set of all points that the solution trajectory from ξ approaches arbitrarily closely infinitely often in the weak topology.¹⁹ The elements of $\omega(\xi)$ are called ω -limit points. The set $\Omega := \bigcup_{\xi \in \mathcal{M}_1^+(S)} \omega(\xi)$ denotes the set of all ω -limit points of all solution trajectories. Then for imitative dynamics in potential game F , we have the following convergence and stability results:²⁰

R1) $\Omega = \text{set of rest points} = RE(F)$;

R2) if F has a unique Nash equilibrium μ^* , then μ^* is Lyapunov stable.

4.2 Cultural Polarization

In this section, we model cultural intolerance as a strictly increasing function of *cultural distance*, that is, the distance between two traits in the trait space. This captures people's general tendencies of evaluating culturally more distant people with stronger biases, which is consistent with the recent findings in the literature of empirical cultural economics, e.g., Guiso, Sapienza and Zingales (2006, 2009), Felbermayr and Toubal (2010), Hahn (2013), and Spolaore and Waacziarg (2009, 2013), who suggest that people have greater intolerance for culturally more distant people. Also, as argued by Akerlof and Kranton (2000), people tend to identify with culturally similar people but to distinguish themselves from culturally distant people.

Consider $\Delta_{zy} = h(|z - y|)$, where $h : [0, 1] \rightarrow [0, 1]$ is a continuous and strictly increasing function with $h(0) = 0$. That is, the cultural intolerance a z -parent has towards trait y depends only on the distance between z and y in the trait space. We call such a function h the *cultural intolerance function*.

In what follows, we investigate how the curvature of the cultural intolerance function h shapes the predictions of the cultural evolutionary dynamic. First, we explore the case in which the cultural intolerance function is convex.

Theorem 2 *Suppose that the cultural intolerance function h is convex. Then the population state $\mu^* = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ is the unique Nash equilibrium of the population game F defined by (2). Thus since F is a potential game, μ^* is Lyapunov stable under the cultural evolutionary dynamic.*

¹⁸Initial condition means initial distribution of strategies (or traits).

¹⁹See the formal definition of $\omega(\xi)$ in Appendix A.1.

²⁰(R1) is the same as Part (a) of Theorem 3 in Cheung (2016). (R2) follows from Lemma 2, Theorem 3 and Remark 5 in Cheung (2014), and Proposition 1 in Cheung (2016).

Proof. See Appendix A.4.

Theorem 2 shows that convexity of intolerance in cultural distance leads to a polarized state in which only the two most extreme traits coexist. The intuition of the proof is as follows. Since the cultural intolerance function is convex, cultural intolerance is intensified as cultural distance increases. Therefore, given any population state, either one (or both) of the two most extreme traits 0 and 1 is associated with the highest aggregate cultural intolerance because these two traits are furthest away from other traits on average. Hence, in the induced population game F defined by (2) where the intolerance function is viewed as the single match payoff function, either one (or both) of 0 and 1 is a best response. As long as there are positive masses on both 0 and 1, the dynamic always has the tendency to move away from the current population state and increases the mass on either one or both of 0 and 1.

Nevertheless, since the cultural evolutionary dynamic is equivalent to an imitative dynamic, if initially the mass on either 0 or 1 is zero, then the dynamic can never lead to a population state with positive masses on both 0 and 1. The intuition here is that when the children search for role models, they will never meet a role model that does not exist at the first place. Therefore, we need to consider some more general long-run predictions of the dynamic that are not necessarily stable.

Note that Ω , the set of all ω -limit points of all solution trajectories, depends on the dynamic studied, which in turn depends on the cultural intolerance function h . From now on, we denote by $\Omega(h)$ the set of all ω -limit points of all solution trajectories for the cultural evolutionary dynamic under cultural intolerance function h .

The following proposition shows that when the cultural intolerance function is convex, the ω -limit points (also rest points) are precisely those population states that consist of no more than two traits, with equal weights on each when the traits are distinct.

Proposition 3 *Suppose that the cultural intolerance function h is convex and that F is the population game defined by (2). Then*

$$\Omega(h) = \text{set of rest points} = RE(F) = \left\{ \frac{1}{2}\delta_a + \frac{1}{2}\delta_b : 0 \leq a \leq b \leq 1 \right\}.$$

Proof. See Appendix A.4.

Note that under imitative dynamics, any non-Nash rest points are not Lyapunov stable. Hence, when the cultural intolerance function is convex, the most extremely polarized state $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ is the only ω -limit point (also rest point) that is Lyapunov stable.

Next, we consider the case in which the cultural intolerance function is strictly concave. Intuitively, any strictly concave cultural intolerance function can be viewed as less extreme than any convex cultural intolerance function. One may conjecture that polarization is more difficult to sustain and more diversified population states tend to appear as ω -limit points. The following theorem confirms this conjecture.

Theorem 3 *Suppose that the cultural intolerance function h is strictly concave. We have the following:*

- i) *Under h , the population state $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ is not a Nash equilibrium of the population game F defined by (2), and thus is not Lyapunov stable under the cultural evolutionary dynamic.*
- ii) *Let \tilde{h} be a convex cultural intolerance function. Then $\Omega(h) \supsetneq \Omega(\tilde{h})$. In particular, $\Omega(h)$ contains some more diversified population states (i.e., states with supports having more than two different traits).*

Proof. See Appendix A.4.

Part (i) of Theorem 3 shows that polarization is no longer a stable phenomenon when the cultural intolerance function is strictly concave. Part (ii) shows that when the cultural intolerance function is strictly concave, the set of ω -limit points (also rest points) for the cultural evolutionary dynamic is richer and contains some more diversified population states. The intuition behind Theorem 3 is that when the increment in cultural intolerance decreases as cultural distance increases, it is no longer the case that either one (or both) of the two most extreme traits 0 and 1 is associated with the highest aggregate cultural intolerance for any population state. Instead, at states where the distribution of traits places most of the weight on 0 and 1, some trait different from the two most extreme traits 0 and 1 is associated with even higher aggregate cultural intolerance compared to 0 and 1. Therefore, the dynamic always has the tendency to move away from these states.

From the above results, we see that whether the cultural intolerance function h is convex or concave is critical for polarization. That h is convex is a sufficient condition for the population state $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ to be the only ω -limit point (also rest point) that is Lyapunov stable. Also, that h is not strictly concave is a necessary condition for $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ to be the only ω -limit point that is Lyapunov stable.

5 Conclusion

This paper generalizes the discrete cultural transmission model proposed by Bisin and Verdier (2001) to continuous trait space. We show that under cultural substitutability, the cultural evolutionary dynamic is equivalent to an imitative dynamic in a population game in which a player's payoff is equal to the aggregate cultural intolerance he has towards other agents. We find that as in Bisin and Verdier (2001), cultural substitutability leads to cultural heterogeneity.

We then model an agent's cultural intolerance towards another agent in the population as a strictly increasing function of cultural distance. This captures people's general tendencies of evaluating culturally more distant people with stronger biases. We find that whether the cultural intolerance function is convex or concave is crucial and leads to distinct cultural phenomena. When the cultural intolerance function is convex, as time passes the population will be divided into two groups with the two most extreme traits in the trait space. Moreover, such a population state is the only ω -limit point that is Lyapunov stable under the cultural evolutionary dynamic. On the other hand, when the cultural intolerance function is strictly concave, polarization is no longer stable and more diversified population states consisting of more traits appear as ω -limit points.

There are various possible directions for future study. First, in our model agents do not interact with each others within their own generation. Extending the current model to allow social interactions would be an important extension and would help to connect the literature of cultural transmission tighter to the literature of evolution.²¹ Second, the trait space we consider in this paper is a closed interval on the real line. It is of interest to investigate whether polarization could arise when the trait space is more complicated (e.g., if the trait space is a unit circle, which can be used to describe people's preferences for geographic locations).

A Appendix

A.1 Weak Topology, and Definitions for the Study of Stability and Convergence

The weak topology is related to weak convergence of measures. The weak topology on $\mathcal{M}_1^+(T)$ is metrized by the Prohorov metric κ , which is defined by

$$\kappa(\mu, \nu) := \inf\{\varepsilon > 0 : \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon, \forall A \in \mathcal{B}\},$$

²¹See Heifetz, Shannon and Spiegel (2007a,b) for their study of evolution on continuous trait spaces.

where $A^\varepsilon := \{x \in T : |x - y| < \varepsilon \text{ for some } y \in A\}$.²² Specifically, $\mu_n \in \mathcal{M}_1^+(T)$ converges weakly to $\mu \in \mathcal{M}_1^+(T)$ if and only if $\kappa(\mu_n, \mu) \rightarrow 0$. In other words, weak convergence and convergence in the Prohorov metric are equivalent. Also, if $\mu = (1 - \varepsilon)\delta_x + \varepsilon\delta_u$ with $0 \leq \varepsilon \leq 1$, where δ is the Dirac delta measure,²³ then $\kappa(\mu, \delta_x) = \min\{\varepsilon, |x - u|\}$. In particular, the distance between two monomorphic states in the weak topology agrees with the underlying metric in the continuous trait space T , i.e., $\kappa(\delta_x, \delta_u) = |x - u|$, when x and u are close to each other in T . We use the Prohorov metric κ to measure the distances between population states.

Let $\mu \in \mathcal{M}_1^+(T)$ and $Y \subseteq \mathcal{M}_1^+(T)$. The distance between μ and the set Y in the weak topology is $\kappa(\mu, Y) := \inf\{\kappa(\mu, \nu) : \nu \in Y\}$. The ε -neighborhood of Y (in the weak topology) is $Y^\varepsilon := \{\mu \in \mathcal{M}_1^+(T) : \kappa(\mu, Y) < \varepsilon\}$. Consider an evolutionary dynamic on $\mathcal{M}_1^+(T)$ (like our cultural evolutionary dynamic) that admits a unique forward solution from each initial condition, and suppose that solutions to the dynamic are continuous in their initial conditions. Let $Z \subseteq \mathcal{M}_1^+(T)$ be a closed set. We say Z is *Lyapunov stable* under the dynamic if for every $\varepsilon > 0$ there exists $\delta > 0$ such that every solution of the dynamic that starts in Z^δ is contained in Z^ε .

Let $\xi \in \mathcal{M}_1^+(T)$, and let $\{\mu_t\}_{t \in [0, \infty)}$ be the solution trajectory to the dynamic with $\mu_0 = \xi$. The ω -limit $\omega(\xi)$ is the set of all points that the solution trajectory from ξ approaches arbitrarily closely infinitely often in the weak topology:

$$\omega(\xi) := \{\psi \in \mathcal{M}_1^+(T) : \exists \{t_k\}_{k=1}^\infty \text{ with } \lim_{k \rightarrow \infty} t_k = \infty \text{ such that } \mu_{t_k} \xrightarrow{w} \psi \text{ as } k \rightarrow \infty\}.$$

Since $\mathcal{M}_1^+(T)$ is compact in the weak topology, $\omega(\xi)$ is nonempty. The set $\Omega := \bigcup_{\xi \in \mathcal{M}_1^+(T)} \omega(\xi)$ denotes the set of all ω -limit points of all solution trajectories, which provides a basic notion of recurrence for deterministic dynamics.

A.2 Proof of Proposition 1

Since

$$\begin{aligned} & d(e_z)V_{zz} + (1 - d(e_z)) \int_{y \in T} V_{zy} \mu(dy) - c(e_z) \\ &= \int_{y \in T} V_{zy} \mu(dy) + d(e_z) \int_{y \in T} (V_{zz} - V_{zy}) \mu(dy) - c(e_z) \\ &= \int_{y \in T} V_{zy} \mu(dy) + d(e_z) \int_{y \in T} \Delta_{zy} \mu(dy) - c(e_z) \end{aligned}$$

²²Cf. Billingsley, 1999, p. 72–73.

²³ δ_x is the monomorphic state in which all agents have trait $x \in T$.

$$= \int_{y \in T} V_{zy} \mu(dy) + d(e_z) \Delta_z(\mu) - c(e_z),$$

the maximization problem (M_z) is equivalent to

$$\max_{e_z} d(e_z) \Delta_z(\mu) - c(e_z). \quad (M')$$

Taking the first order condition, we have

$$d'(e_z) \Delta_z(\mu) = c'(e_z),$$

which, writing $(\frac{c'}{d'})^{-1}$ for $\frac{(c')^{-1}}{(d')^{-1}}$, gives

$$e_z^* = \left(\frac{c'}{d'}\right)^{-1}(\Delta_z(\mu)). \quad (5)$$

Note that since c' is strictly increasing and d' is weakly decreasing, the function $\frac{c'}{d'}$ is strictly increasing, and hence its inverse $(\frac{c'}{d'})^{-1}$ exists and is strictly increasing. From the assumptions on $c(\cdot)$ and $d(\cdot)$, we have

$$d''(e_z) \Delta_z(\mu) - c''(e_z) < 0.$$

So the objective function in (M') is strictly concave in e_z , and hence e_z^* is the unique maximizer. Since $c'(0) = 0$, we have

$$\frac{c'(0)}{d'(0)} = 0 \leq \Delta_z(\mu).$$

Applying $(\frac{c'}{d'})^{-1}$ yields

$$0 \leq \left(\frac{c'}{d'}\right)^{-1}(\Delta_z(\mu)) = e_z^*.$$

Since $\Delta_z(\mu) \in [0, 1]$ and $(\frac{c'}{d'})^{-1}$ is continuous, $e_z^* \geq 0$ is bounded.

Finally, since $(\frac{c'}{d'})^{-1}$ is a strictly increasing function, e_z^* strictly increases with $\Delta_z(\mu)$, the aggregate cultural intolerance of a z -parent. *Q.E.D.*

A.3 Some proofs from Section 3

Proof of Proposition 2. Let $z \in T$ and consider monomorphic state δ_z . Recall that $F_x(\mu) := \int_{y \in T} \Delta_{xy} \mu(dy)$. So $F_x(\delta_z) = \Delta_{xz}$. Then $F_z(\delta_z) = 0 < F_y(\delta_z)$ for any $y \in T \setminus \{z\}$. Hence, δ_z is not a Nash equilibrium. *Q.E.D.*

A.4 Some proofs from Section 4

Proof of Theorem 2. First, we prove that μ^* is a Nash equilibrium. For any $x \in (0, 1)$,

$$\begin{aligned} F_x(\mu^*) &= \frac{1}{2}(h(x) + h(1-x)) \\ &= \frac{1}{2}(h(x \cdot 1 + (1-x) \cdot 0) + h((1-x) \cdot 1 + x \cdot 0)) \\ &\leq \frac{1}{2}(xh(1) + (1-x)h(0) + (1-x)h(1) + xh(0)) \text{ (by convexity of } h(\cdot)\text{)} \\ &= \frac{1}{2}h(1) = F_0(\mu^*) = F_1(\mu^*). \end{aligned}$$

Therefore, μ^* is a Nash equilibrium.

Next we prove that the Nash equilibrium is unique. To do so, it suffices to prove that for any $\mu \neq \mu^*$ and any $x \in (0, 1)$, x is never a best response. Let us write down the payoffs corresponding to all the strategies for any $\mu \in \mathcal{M}_1^+(T)$:

$$\begin{aligned} F_0(\mu) &= \int_{[0,1]} h(y) \mu(dy); \\ F_1(\mu) &= \int_{[0,1]} h(1-y) \mu(dy); \end{aligned}$$

and for any $x \in (0, 1)$,

$$F_x(\mu) = \int_{[0,x]} h(x-y) \mu(dy) + \int_{[x,1]} h(y-x) \mu(dy).$$

Suppose for a $x \in (0, 1)$ such that $F_x(\mu) \geq F_0(\mu)$, which implies

$$\begin{aligned} F_x(\mu) - F_0(\mu) &= \int_{[0,x]} h(x-y) \mu(dy) + \int_{[x,1]} h(y-x) \mu(dy) - \int_{[0,1]} h(y) \mu(dy) \\ &= \int_{[0,x]} (h(x-y) - h(y)) \mu(dy) + \int_{[x,1]} (h(y-x) - h(y)) \mu(dy) \\ &= \int_{[0,x]} (h(x) + h(x-y) - h(y) - h(x)) \mu(dy) \\ &\quad + \int_{[x,1]} (-h(x) + h(y-x) - h(y) + h(x)) \mu(dy) \end{aligned}$$

$$\begin{aligned}
&= h(x)(\mu([0, x]) - \mu([x, 1])) + \int_{[0, x]} (h(x - y) - h(y) - h(x)) \mu(dy) \\
&\quad + \int_{[x, 1]} (h(y - x) - h(y) + h(x)) \mu(dy) \geq 0.
\end{aligned}$$

Since, $h(x - y) - h(y) - h(x) \leq 0$ for any $y \in [0, x]$, we have $\int_{[0, x]} (h(x - y) - h(y) - h(x)) \mu(dy) \leq 0$. Moreover, by convexity of $h(\cdot)$, $h(y - x) + h(x) \leq h(y)$ for any $y \in [x, 1]$. Hence, $\int_{[x, 1]} (h(y - x) - h(y) + h(x)) \mu(dy) \leq 0$. This implies that $\mu([0, x]) - \mu([x, 1]) \geq 0$. Let $\mu([0, x]) - \mu([x, 1]) = \sigma$, where $\sigma \geq 0$. Since $\mu([0, x]) + \mu([x, 1]) = 1$, we can rewrite $\mu([0, x]) = \frac{1+\sigma}{2}$ and $\mu([x, 1]) = \frac{1-\sigma}{2}$. On the other hand, we have

$$\begin{aligned}
F_x(\mu) - F_1(\mu) &= \int_{[0, x]} h(x - y) \mu(dy) + \int_{[x, 1]} h(y - x) \mu(dy) - \int_{[0, 1]} h(1 - y) \mu(dy) \\
&= \int_{[0, x]} (h(x - y) - h(1 - y)) \mu(dy) + \int_{[x, 1]} (h(y - x) - h(1 - y)) \mu(dy) \\
&= \int_{[0, x]} (h(x) + h(x - y) - h(1 - y) - h(x)) \mu(dy) \\
&\quad + \int_{[x, 1]} (-h(x) + h(y - x) - h(1 - y) + h(x)) \mu(dy) \\
&= h(x)(\mu([0, x]) - \mu([x, 1])) + \int_{[0, x]} (h(x - y) - h(1 - y) - h(x)) \mu(dy) \\
&\quad + \int_{[x, 1]} (h(y - x) - h(1 - y) + h(x)) \mu(dy).
\end{aligned}$$

Given that $\mu([x, 1]) = \frac{1-\sigma}{2}$, then

$$\int_{[x, 1]} (h(y - x) - h(1 - y) + h(x)) \mu(dy) \leq \frac{1-\sigma}{2} (h(1 - x) + h(x)),$$

and equality holds if and only if $\mu(\{1\}) = \frac{1-\sigma}{2}$ since $h(\cdot)$ is strictly increasing. Moreover, by convexity of $h(\cdot)$, $h(x - y) + h(1 - x) \leq h(1 - y)$ for any $y \in [0, x]$. So,

$$\begin{aligned}
\int_{[0, x]} (h(x - y) - h(1 - y) - h(x)) \mu(dy) &\leq \int_{[0, x]} (-h(1 - x) - h(x)) \mu(dy) \\
&= -\mu([0, x]) (h(1 - x) + h(x)) \\
&= -\frac{1+\sigma}{2} (h(1 - x) + h(x)). \tag{6}
\end{aligned}$$

Hence, we have

$$\begin{aligned}
F_x(\mu) - F_1(\mu) &\leq h(x)\sigma - \frac{1+\sigma}{2} (h(1 - x) + h(x)) + \frac{1-\sigma}{2} (h(1 - x) + h(x)) \\
&= -h(1 - x)\sigma
\end{aligned}$$

$$\leq 0,$$

where both equalities hold if and only if $\sigma = 0$ and $\mu([x, 1]) = \mu(\{1\}) = \frac{1}{2}$ and equality holds in (6). Note that when $\sigma = 0$ and $\mu([x, 1]) = \mu(\{1\}) = \frac{1}{2}$,

$$F_x(\mu) - F_0(\mu) = \int_{[0,x)} (h(x-y) - h(y) - h(x)) \mu(dy) + \frac{1}{2}(h(1-x) - h(1) + h(x)).$$

Since we assume that $F_x(\mu) - F_0(\mu) \geq 0$ and by convexity of $h(\cdot)$, $\frac{1}{2}(h(1-x) - h(1) + h(x)) \leq 0$, we need $\int_{[0,x)} (h(x-y) - h(y) - h(x)) \mu(dy) \geq 0$. However, this can only happen when $\mu([0, x)) = \mu(\{0\}) = \frac{1}{2}$.

Therefore, for any $\mu \neq \mu^*$, if a $x \in (0, 1)$ yields a weakly higher payoff than 0 does, i.e., $F_x(\mu) \geq F_0(\mu)$, then it must be the case that 1 yields a strictly higher payoff than x does, i.e., $F_1(\mu) > F_x(\mu)$. Similarly, we can prove that if a $x \in (0, 1)$ yields a weakly higher payoff than 1 does, i.e., $F_x(\mu) \geq F_1(\mu)$, then it must be the case that 0 yields a strictly higher payoff than x does, i.e., $F_0(\mu) > F_x(\mu)$. This proves that $x \in (0, 1)$ is never a best response for any $\mu \neq \mu^*$. In other words, in any Nash equilibrium μ of the population game, $\mu((0, 1)) = 0$.

It is straightforward to show that for any $\mu \in \mathcal{M}_1^+(T)$ that satisfies $\mu((0, 1)) = 0$, if $\mu(\{0\}) > \mu(\{1\})$, then $F_1(\mu) > F_0(\mu)$, and if $\mu(\{1\}) > \mu(\{0\})$, then $F_0(\mu) > F_1(\mu)$. Therefore, μ^* is the unique Nash equilibrium. *Q.E.D.*

Proof of Proposition 3. It is trivial that monomorphic states are restricted equilibria. Any dimorphic state having the form $\mu = \frac{1}{2}\delta_x + \frac{1}{2}\delta_y$ ($x \neq y$) satisfies $F_x(\mu) = \frac{1}{2}h(|x-y|) = F_y(\mu)$, and so is a restricted equilibrium. In addition, any dimorphic state $\mu = \alpha\delta_x + (1-\alpha)\delta_y$ with $\alpha \neq \frac{1}{2}$ is not a restricted equilibrium since $F_x(\mu) \neq F_y(\mu)$.

Next we prove that any population state μ having more than 2 points in its support cannot be a restricted equilibrium. Let x be the minimum of $\text{supp}(\mu)$ and y be the maximum of $\text{supp}(\mu)$.²⁴ Suppose there exists $z \in \text{supp}(\mu)$ such that $z \neq x$ and $z \neq y$. Then by the same logic as in the proof of Theorem 2, we have either $F_x(\mu) > F_z(\mu)$ or $F_y(\mu) > F_z(\mu)$. Hence, μ is not a restricted equilibrium. *Q.E.D.*

Proof of Theorem 3.

Part (i): Let $\tilde{\mu} = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. Then $F_0(\tilde{\mu}) = \frac{1}{2}h(1) = F_1(\tilde{\mu})$, but

$$F_{\frac{1}{2}}(\tilde{\mu}) = \frac{1}{2}h\left(\frac{1}{2}\right) + \frac{1}{2}h\left(\frac{1}{2}\right) = h\left(\frac{1}{2}\right) = h\left(\frac{1}{2} \times 1 + \frac{1}{2} \times 0\right) > \frac{1}{2}h(1) + \frac{1}{2}h(0) = \frac{1}{2}h(1).$$

²⁴Note that $\text{supp}(\mu)$ is a closed set.

Hence, $\tilde{\mu}$ is not a Nash equilibrium.

Part (ii): Any $\mu = \frac{1}{2}\delta_a + \frac{1}{2}\delta_b$ with $0 \leq a \leq b \leq 1$ is a restricted equilibrium when $h(\cdot)$ is strictly concave. Moreover, it is easy to check that $\frac{\frac{1}{2}h(\frac{1}{2})}{2h(\frac{1}{2})-\frac{1}{2}h(1)}\delta_0 + \frac{h(\frac{1}{2})-\frac{1}{2}h(1)}{2h(\frac{1}{2})-\frac{1}{2}h(1)}\delta_{\frac{1}{2}} + \frac{\frac{1}{2}h(\frac{1}{2})}{2h(\frac{1}{2})-\frac{1}{2}h(1)}\delta_1$ is a restricted equilibrium when $h(\cdot)$ is strictly concave. *Q.E.D.*

References

- Akerlof, G.A., and Kranton, R.E., 2000. Economics and identity. *Quarterly Journal of Economics* 115, 715–753.
- Bisin, A., and Topa, G., 2003. Empirical models of cultural transmission. *Journal of the European Economic Association* 1, 363–375.
- Bisin, A., Topa, G., and Verdier, T., 2004. Religious intermarriage and socialization in the United States. *Journal of Political Economy* 112, 615–664.
- Bisin, A., Topa, G., and Verdier, T., 2009. Cultural transmission, socialization and the population dynamics of multiple-trait distributions. *International Journal of Economic Theory* 5, 139–154.
- Bisin, A., and Verdier, T., 2001. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory* 97, 298–319.
- Bisin, A., and Verdier, T., 2010. The economics of cultural transmission and socialization, in: Benhabib, J., Bisin, A., and Jackson, M.O. (Eds.), *Handbook of Social Economics*. Elsevier, Amsterdam.
- Boyd, R., and Richerson, P.J., 1985. *Culture and the Evolutionary Process*. University of Chicago Press, Chicago.
- Buechel, B., Hellmann, T., and Pichler, M.M., 2014. The dynamics of continuous cultural traits in social networks. *Journal of Economic Theory* 154, 274–309.
- Cavalli-Sforza, L.L., and Feldman, M.W., 1973. Cultural versus biological inheritance: phenotypic transmission from parents to children. *American Journal of Human Genetics* 25, 618–637.
- Cavalli-Sforza, L.L., and Feldman, M.W., 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press, Princeton.
- Cheung, M.-W., 2014. Pairwise comparison dynamics for games with continuous strategy space. *Journal of Economic Theory* 153, 344–375.
- Cheung, M.-W., 2016. Imitative dynamics for games with continuous strategy space. *Games and Economic Behavior* 99, 206–223.

- Dixit, A.K., and Weibull, J.W., 2007. Political polarization. *Proceedings of National Academy of Science* 104, 7351–7356.
- Doepke, M., and Zilibotti, F., 2008. Occupational choice and the spirit of capitalism. *Quarterly Journal of Economics* 123, 747–793.
- Downs, A., 1957. *An Economic Theory of Democracy*. Harper, New York.
- Felbermayr, G.J., and Toubal, F., 2010. Cultural proximity and trade. *European Economic Review* 54, 279–293.
- Guiso, L., Sapienza, P., and Zingales, L., 2006. Does culture affect economic outcomes? *Journal of Economic Perspectives* 20, 23–48.
- Guiso, L., Sapienza, P., and Zingales, L., 2009. Cultural biases in economic exchange? *Quarterly Journal of Economics* 124, 1095–1131.
- Hahn, F.R., 2014. Culture, geography and institutions: empirical evidence from small-scale banking. *Economic Journal* 124, 859–886.
- Hauk, E., and Saez-Marti, M., 2002. On the cultural transmission of corruption. *Journal of Economic Theory* 107, 311–335.
- Heifetz, A., Shannon, C., and Spiegel, Y., 2007a. The dynamic evolution of preferences. *Economic Theory* 32, 251–286.
- Heifetz, A., Shannon, C., and Spiegel, Y., 2007b. What to maximize if you must. *Journal of Economic Theory* 133, 31–57.
- Hotelling, H., 1929. Stability in competition. *Economic Journal* 39, 41–57.
- Hunter, J.D., 1992. *Cultural Wars: The Struggle to Define America*. Basic Books, New York.
- McMurray, J., 2015. Ideology as opinion: polarization in a spatial model of common-value elections. Working paper.
- Montgomery, J.D., 2010. Intergenerational cultural transmission as an evolutionary game. *American Economic Journal: Microeconomics* 2, 115–136.
- Nickerson, R.S., 1998. Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology* 2, 175–220.
- Oechssler, J., and Riedel, F., 2001. Evolutionary dynamics on infinite strategy spaces. *Economic Theory* 17, 141–162.
- Oechssler, J., and Riedel, F., 2002. On the dynamic foundation of evolutionary stability in continuous models. *Journal of Economic Theory* 107, 223–252.
- Panbianco, F., 2014. Socialization networks and the transmission of interethnic attitudes. *Journal of Economic Theory* 150, 583–610.

- Pichler, M.M., 2010. The economics of cultural formation of preferences. Working Paper 431, Institute of Mathematical Economics, Bielefeld University.
- Rabin, M., and Schrag, J.L., 1999. First impressions matter: a model of confirmatory bias. *Quarterly Journal of Economics* 114, 37–82.
- Rudin, W., 1987. *Real and Complex Analysis*, 3rd edition. McGraw-Hill, New York.
- Sandholm, W.H., 2010. *Population Games and Evolutionary Dynamics*. MIT Press, Cambridge.
- Sartori, G., 1966. European political parties: the case of polarized pluralism, in: LaPalombara, J., and Weiner, M. (Eds.), *Political Parties and Political Development*. Princeton University Press, Princeton.
- Sartori, G., 1976. *Parties and Party Systems: A Framework for Analysis*. Cambridge University Press, New York.
- Spolaore, E., and Wacziarg, R., 2009. The diffusion of development. *Quarterly Journal of Economics* 124, 469–529.
- Spolaore, E., and Wacziarg, R., 2013. How deep are the roots of economic development? *Journal of Economic Literature* 51, 325–369.
- Tabellini, G., 2008. The scope of cooperation: values and incentives. *Quarterly Journal of Economics* 123, 905–950.
- Vaughan, D., 2010. Conformity and the transmission of continuous cultural traits. Departmental Working Paper 2011-01, Department of Economics, New York University.