working paper

17 20

Convergence of Cultural Traits with Time-Varying Self-Confidence in the Panebianco (2014) Model—A Corrigendum

Fabrizio Panebianco, Anja Prummer, and Jan-Peter Siedlarek

FEDERAL RESERVE BANK OF CLEVELAND

ISSN: 2573-7953

Working Paper 17-20 November 2017

Convergence of Cultural Traits with Time-Varying Self-Confidence in the Panebianco (2014) Model—A Corrigendum

Fabrizio Panebianco, Anja Prummer, and Jan-Peter Siedlarek

We highlight that convergence in repeated averaging models commonly used to study cultural traits or opinion dynamics is not equivalent to convergence in Markov chain settings if transition matrices are time-varying. We then establish a new proof for the convergence of cultural traits in the model of Panebianco (2014) correcting the existing proof. The new proof provides novel insights on the long-run outcomes for inessential individuals. We close with a discussion of conditions for convergence in repeated averaging models with time-varying transition matrices.

JEL codes: D83, D85, Z13.
Keywords: Cultural transmission, continuous cultural traits, social networks, opinion dynamics.

Suggested citation: Panebianco, Fabrizio, Anja Prummer, and Jan-Peter Siedlarek, 2017. "Convergence of Cultural Traits with Time-Varying Self-Confidence in the Panebianco (2014) Model—A Corrigendum," Federal Reserve Bank of Cleveland, Working Paper no. 17-20. https://doi.org/10.26509/frbc-wp-201720.

Fabrizio Panebianco is at Bocconi University and IGIER (fabrizio.panebianco@unibocconi.it), Anja Prummer is at School of Economics and Finance, Queen Mary University of London (a.prummer@qmul.ac.uk), and Jan-Peter Siedlarek is at the Federal Reserve Bank of Cleveland (jan-peter.siedlarek@clev.frb.org).

1 Introduction

Panebianco (2014) (from here on "P14") studies a model of continuous trait transmission with inter-ethnic attitudes through parental (vertical) and non-parental (oblique) socialization. P14 establishes convergence of cultural traits and studies the structure of steady state outcomes. This note demonstrates that the proof of the convergence results in P14 is incorrect in two places. In the first instance, the proof in P14 incorrectly transposes a convergence result for Markov Chains which does not apply to the model in P14. In the second instance, an algebraic argument in P14 contains a mistake. The note provides a new proof that corrects both issues and recovers all affected results of P14.

We first highlight the difference between Markov chains and repeated averaging models, which are commonly used in models of cultural transmission and opinion formation, in a simple example in Section 2. In Section 3 we identify the two errors in the convergence proof of P14 and present a new proof that restores all results and offers some novel insights into the steady state properties of the P14 model. We conclude with a brief discussion of convergence in repeated averaging models with time-varying transition matrices of which P14 is an example.

2 Example

Before we present the correction to P14 in detail, we briefly illustrate the repeated averaging setting and its relationship with Markov chains in a simple example. Fix a sequence of row stochastic matrices $\{\mathbf{X}_t\}$ that is time-varying by alternating between the two matrices $\mathbf{X}_{\text{odd}}$ and $\mathbf{X}_{\text{even}}$ depending on whether the period t is odd or even.

$$\mathbf{X}_{\text{odd}} = \begin{pmatrix} 0.5 & 0.1 & 0.4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{X}_{\text{even}} = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Define by $\mathbf{X}_t^{\text{Right}}$ and $\mathbf{X}_t^{\text{Left}}$ the products resulting from multiplying on the right and left, respectively, a total number of t matrices according to the sequence of $\mathbf{X}_{\text{odd}}$ and $\mathbf{X}_{\text{even}}$, starting with $\mathbf{X}_{\text{odd}}$.

Multiplication on the right as in $\mathbf{X}_t^{\text{Right}}$ presents a Markov chain. In a Markov chain the dimensions of $\mathbf{X}_t$ correspond to states and $\mathbf{X}_t$ is a transition matrix in which element $x_{ij}(t)$ describes the probability of transitioning from state i to j. The right product converges towards the matrix

$$\begin{aligned}\lim_{t\to\infty}\mathbf{X}_t^{\text{Right}} =&\mathbf{X}_{\text{odd}}\mathbf{X}_{\text{even}}\mathbf{X}_{\text{odd}}\mathbf{X}_{\text{even}}\dots\\ =&\begin{pmatrix}0 & 0.4 & 0.6\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.\end{aligned}$$

Note that the long-run outcome depends on the first matrix on the left of the sequence. If the sequence started with $\mathbf{X}_{\text{even}}$ the positions of 0.4 and 0.6 would be switched around in the long-run outcome.

Multiplication on the left represents a repeated averaging setting that is used in the cultural traits model of P14. This type of model is also used in naive learning and opinion formation literature, including for example, Cavalli-Sforza and Feldman (1973), DeGroot (1974) and, more recently, DeMarzo, Vayanos, and Zwiebel (2003), Golub and Matthew O. Jackson (2010) and Büchel, Hellmann, and Pichler (2014). Here the transition matrix $\mathbf{X}_t$ acts as an influence matrix that describes how next period attitudes are derived as the weighted average of current-period attitudes with element $x_{ij}(t)$ giving the weight that individual i assigns to the trait of individual j.

In contrast to the Markov chain approach above, $\mathbf{X}_t^{\text{Left}}$ does not converge but instead leads to a limit cycle which alternates between two matrices depending on whether the final matrix on the left is $\mathbf{X}_{\text{odd}}$ or $\mathbf{X}_{\text{even}}$.

$$\begin{aligned}\lim_{t\to\infty}\mathbf{X}_t^{\text{Left}} &= \dots\mathbf{X}_{\text{even}}\mathbf{X}_{\text{odd}}\mathbf{X}_{\text{even}}\mathbf{X}_{\text{odd}}\\ &=\begin{cases}\begin{pmatrix}0 & 0.4 & 0.6\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} & \text{if } t \text{ is odd, and}\\ \begin{pmatrix}0 & 0.6 & 0.4\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} & \text{if } t \text{ is even.}\end{cases}\end{aligned}$$

Figure 1 illustrates these dynamics and plots the entry in the first row and the third column of $\mathbf{X}_t^{\text{Right}}$ and $\mathbf{X}_t^{\text{Left}}$. In the cultural traits setting, this entry corresponds to the trait held by the first agent if we set the initial trait vector to $\overline{V}_0 = \begin{pmatrix}0 & 0 & 1\end{pmatrix}'$.

The example illustrates that the convergence behaviour of a given sequence of row stochas-

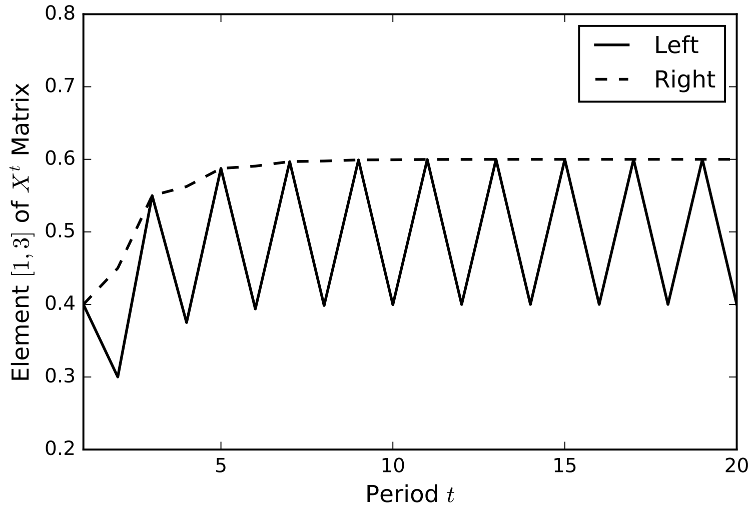

Figure 1: Convergence in Markov Chain and Cultural Transmission Models – Counterexample

tic matrices depends on whether multiplication is from the right as in Markov chain models or the left as in the cultural traits transmission literature discussed here. Furthermore, for left multiplication, the example is a simple sequence of matrices that does not converge in a cultural transmission context. Both points play an important role in the inaccuracies in the convergence proof of P14 that we describe below.

3 Convergence in the Panebianco (2014) Model

The model in P14 presents a model of cultural transmission that can be summarized as follows:

$$\begin{aligned}\overline{V}_{t+1} &= \mathbf{X}_t \overline{V}_t \\ &= \mathbf{X}_t \mathbf{X}_{t-1} \ldots \mathbf{X}_0 \overline{V}_0, \\ &= \mathbf{X}^t_{\text{Left}} \overline{V}_0\end{aligned} \tag{1}$$

where $\overline{V}$ is a column vector of inter-ethnic attitudes and $\mathbf{X}_t$ is a time-varying row stochastic square transition matrix. P14 endows $\mathbf{X}_t$ with the following specific structure:

$$\mathbf{X}_t = \mathbf{S}_t + (\mathbf{I} - \mathbf{S}_t)\,\mathbf{\Phi}, \tag{2}$$

where $\mathbf{S}_t$ is a diagonal square matrix capturing the vertical aspect of socialization within a group and $\mathbf{\Phi}$ is a row stochastic square matrix with entries ϕ_{ij} that captures the oblique socialization between groups. Assumption 1 in Panebianco (2014) ensures that the entries of $\mathbf{S}_t$ and thus the diagonal entries of $\mathbf{X}_t$ are non-zero for all time periods t. For off-diagonal entries this structure implies that the pattern of zeros in $\mathbf{X}_t$ is equal to that of $\mathbf{\Phi}$. Furthermore, the ratio of any pair of off-diagonal entries of $\mathbf{X}_t$ in the same row, that is, the relative socialization weight that any group puts on a given pair of other groups, is constant for all t.

P14 presents one main convergence result for this system and a corollary that presents a generalization of the result to time-varying $\mathbf{\Phi}_t$ under the condition that $\mathbf{\Phi}_t$ has at most one communication class per component and a pattern of zero entries that is constant across time.

A component in a repeated averaging setting refers to a group of individuals such that there is a non-zero weight between every pair in the group in at least one direction after a certain minimum number of periods. This positive weight corresponds to the notion of a directed path through a network if one treats the individuals as a set of vertices and $\mathbf{X}_t$ as an adjacency matrix in which $x_{i,j}(t) > 0$ implies that individual i is influenced by individual j. A communication class then refers to a group of individuals that influence each other but put zero weight on individuals outside the group. Such a class is also referred to as an essential class. Furthermore, individuals that themselves influence every other individual that they are influenced by are called essential. Those that are not essential are called inessential. The convergence results in P14 are as follows.

(P14) Proposition 2 The system described by Equations (1) and (2) converges for any time-invariant row stochastic matrix $\mathbf{\Phi}$.

(P14) Corollary 1 The convergence result can be extended to time-varying $\mathbf{\Phi}_t$ if each $\mathbf{\Phi}_t$ has at most one communication class per component and the zero entries of $\mathbf{\Phi}_t$ are fixed for all $t > T$ for some period T.

The proof of Proposition 2 in P14 distinguishes between transition matrices that are *irreducible* and those that are *reducible*. Irreducible matrices correspond to influence networks that are strongly connected such that every pair of individuals, either directly or indirectly, influences each other. All individuals thus form a single essential class. By contrast, in a reducible transition matrix, there exist some inessential individuals that are influenced by some other individuals that they themselves do not influence. If a matrix is reducible, it can be written in lower triangular block form as illustrated in Equation (3) where those groups of individuals that are essential are collected in the block matrix $\mathbf{X}_{t,[1,1]}$ and the second row collects the remaining inessential individuals.

$$\mathbf{X}_t = \begin{pmatrix} \mathbf{X}_{t,[1,1]} & \mathbf{0} \\ \mathbf{X}_{t,[2,1]} & \mathbf{X}_{t,[2,2]} \end{pmatrix} \tag{3}$$

Reducible matrices are then further subdivided according to whether the individuals in block $\mathbf{X}_{t,[1,1]}$ form one diagonal block and thus a single essential class (Case 1) or more than one diagonal block and thus more than one essential class (Case 2). The convergence proof in P14 is incorrect in its argument for convergence of reducible matrices for both cases.

3.1 Case 1 – One Essential Class

In the proof of Case 1, P14 restates Theorem 3.2 from D'Amico, Janssen, and Manca (2009) which establishes convergence of single-unireducible non-homogeneous Markov chains, and then builds on this result. However, in restating it as Theorem 2 on p.602, P14 switches the direction of multiplication from the right as in the original to the left as needed for the model in P14. It thus incorrectly applies a result from Markov chains to a repeated average setting. The two classes of models show different convergence behaviours for time-varying matrices as we show in the example in Section 2.

Convergence for the case of reducible matrices with a single diagonal block can be readily recovered by using an appropriate convergence result for left multiplication. Theorem 1.10 in Hartfiel (2006) provides this result. The theorem relies on the notion of *regular* matrices. A stochastic matrix $\mathbf{A}$ is *regular* if it has exactly one essential class and the upper left block in lower triangular form is primitive, that is, there exists a constant k such that $\mathbf{A}_{1,1}^k$ has all strictly positive entries.

(Hartfiel (2006), Theorem 1.10) If $\mathbf{B}_{p,h}$ is regular for each $p \geq 0, h > 1$ and

$$\min_{i,j}{}^{+} a_{ij}(k) \geq \gamma > 0$$

uniformly for all $k \geq 1$ (where $\min^{+}$ is the minimum over all positive entries), then $\lim_{h\to\infty} \mathbf{B}_{p,h} = \mathbf{Y}$, a rank one matrix that depends on p. Further there are constants K and β, $0 < \beta < 1$, such that

$$\left\|\mathbf{B}_{p,h} - \mathbf{Y}\right\| \leq K\beta^{h}.$$

$\mathbf{B}_{p,h}$ denotes the backward product of a sequence of matrices $\mathbf{A}_t$ with elements $a_{ij}(t)$ and is defined as

$$\mathbf{B}_{p,h} = \mathbf{A}_{p+h}\mathbf{A}_{p+h-1}\ldots\mathbf{A}_{p+1}.$$

Proof of Convergence for Case 1. Case 1 of the P14 model satisfies the conditions of Theorem 1.10

Hartfiel (2006). First, the left multiplication in Equation (1) corresponds to backward products. Second, in Case 1 there is exactly one essential or communication class. Third, the block matrix corresponding to this class is primitive. By the definition of a communication class there exists a path from every individual within the class to every other individual within the class. Furthermore, as $\mathbf{X}_t$ has non-zero diagonal entries, the block matrix is aperiodic and thus there exists an integer $k > 0$ such that all entries of the k-step backward product of $\mathbf{X}_{t,[1,1]}$ are positive. Finally, Assumption 1 in P14 together with fixed oblique socialization matrix Φ ensures that the non-zero entries of the transition matrix $\mathbf{X}_t$ are bounded away from zero. The left product is thus regular for all t.

It follows from Theorem 1.10 Hartfiel (2006) that $\mathbf{X}_t^{\text{Left}}$ converges and the long-run outcome $\overline{\mathbf{X}}$ is a matrix of rank one, implying consensus in cultural traits. Furthermore, there are constants K and $\beta_X < 1$ such that $\|\mathbf{X}_t^{\text{Left}} - \overline{\mathbf{X}}\| \leq K\beta_X^t$ □

The proof extends to the case of time-varying Φ_t covered by Corollary 1. The additional condition of the corollary that the pattern of zeros remains constant ensures that the regularity of the left product is preserved. Thus as long as the non-zero elements in Φ_t are bounded away from zero for all t, Hartfiel (2006) Theorem 1.10 continues to apply and the cultural traits converge to consensus.

3.2 Case 2 – More Than One Essential Class

To show convergence for Case 2, that is, a reducible transition matrix with more than one isolated block matrix in $\mathbf{X}_{t,[1,1]}$, P14 presents a proof by construction. The proof decomposes each updating step of an individual trait into a weighted average of the previous value of the trait and the long-run outcomes of the essential classes.[1] The argument in P14 is incorrect because the terms that describe the weight assigned to the long-run outcomes of the essential classes *do not* converge with arbitrary time-varying entries in the transition matrix. Specifically, the assertion that "$\sum_{i=1}^{t} \beta_i \frac{\alpha_t!}{\alpha_i!}$ *is a monotone increasing [in t] series*" in P14 (top of p. 605) is not true without further restrictions.

To see why, rewrite this sum term by defining $b(t) \equiv \sum_{i=1}^{t} \beta_i \frac{\alpha_t!}{\alpha_i!}$ and then rearrange as

[1] See the derivation on the bottom half of p.604 of P14.

follows:

$$
\begin{aligned}
b(t) &= \sum_{i=1}^{t} \beta_i \frac{\alpha_t!}{\alpha_i!} \\
&= \alpha_t! \sum_{i=1}^{t} \frac{\beta_i}{\alpha_i!} \\
&= \alpha_t \cdot \alpha_{t-1}! \left[\sum_{i=1}^{t-1} \frac{\beta_i}{\alpha_i!} + \frac{\beta_t}{\alpha_t!} \right] \\
&= \alpha_t b(t-1) + \beta_t
\end{aligned}
$$

$b(t)$ is strictly monotone increasing in t if and only if it is strictly larger than $b(t-1)$. Given that $\alpha_t < 1$, $b(t)$ can be smaller than $b(t-1)$ if β_t is small. In a model with a general time-varying transition matrix $b(t)$ can increase as well as decrease for large t if the individual β_t switch between large and small values.

Note that the proof in P14 does not make use of the specific restrictions on $\mathbf{X}_t$ in that paper and summarized in Equation (2) above. If the proof were valid it would thus apply to a very general class of time-varying transition matrices, including those with time-varying ratios of off-diagonal elements and including the example presented above. As we have shown, this general claim is not true. However, as we argue next, convergence in the model of P14 is preserved, and can be proven by using the restrictions on time variation in the P14 setting.

Proof of Case 2. As $\mathbf{X}_t$ is reducible, the left product $\mathbf{X}_t^{\text{Left}}$ is also reducible and can be written in lower triangular form

$$
\mathbf{X}_t^{\text{Left}} = \begin{pmatrix} \mathbf{X}_{t,[1,1]}^{\text{Left}} & \mathbf{0} \\ \mathbf{X}_{t,[2,1]}^{\text{Left}} & \mathbf{X}_{t,[2,2]}^{\text{Left}} \end{pmatrix}.
$$

We discuss convergence of each block in sequence.

3.2.1 $\mathbf{X}_{t,[1,1]}^{\text{Left}}$ converges to a matrix of rank one $\overline{\mathbf{X}}_{[1,1]}$

Note that for the case of more than one communication class, the block matrix $\mathbf{X}_{t,[1,1]}^{\text{Left}}$ is no longer regular as it contains two communication classes that do not influence each other. Thus, Theorem 1.10 Hartfiel (2006) does not directly imply convergence of $\mathbf{X}_t^{\text{Left}}$ as a whole. However, as the different communication classes present independent blocks consisting of exactly one communication class each, the argument of Case 1 continues to apply to each of them separately. Theorem 1.10 Hartfiel (2006) thus implies convergence of the block matrix $\mathbf{X}_{t,[1,1]}^{\text{Left}}$ to a rank one matrix $\overline{\mathbf{X}}_{[1,1]}$.

3.2.2 $\mathbf{X}^{\mathbf{Left}}_{t,[2,2]}$ converges to the zero matrix

Note first that for every t the lower right block of the transmission matrix $\mathbf{X}_{t,[2,2]}$ is substochastic, that is, the elements in every row sum to less than or equal to one and at least one of the row sums to strictly less than one. Furthermore, as all individuals within this block are inessential, they are each directly or indirectly influenced by an individual outside the block. Therefore there exists a k such that the left product of k times $\mathbf{X}_{t,[2,2]}$ is a substochastic matrix in which all rows add up to strictly less than one. It then follows that $\mathbf{X}^{\mathrm{Left}}_{t,[2,2]} = \mathbf{X}_{t,[2,2]}\mathbf{X}_{t-1,[2,2]} \dots \mathbf{X}_{0,[2,2]}$ is the product of strictly substochastic matrices and converges to zero.

3.2.3 $\mathbf{X}^{\mathbf{Left}}_{t,[2,2]}$ converges to a matrix $\overline{\mathbf{X}}_{[2,2]}$

Finally, we show that the block $\mathbf{X}^{\mathrm{Left}}_{t,[2,1]}$ converges. In P14 a reducible transition matrix $\mathbf{X}_t$ derives from $\mathbf{\Phi}$ being reducible and it can thus be decomposed as follows:

$$
\begin{aligned}
\mathbf{X}_t =& \mathbf{S}_t + (\mathbf{I} - \mathbf{S}_t)\,\mathbf{\Phi} \\
=& \begin{pmatrix} \mathbf{S}_{t,[1,1]} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{t,[2,2]} \end{pmatrix} + \begin{pmatrix} \mathbf{I} - \mathbf{S}_{t,[1,1]} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - \mathbf{S}_{t,[2,2]} \end{pmatrix} \begin{pmatrix} \mathbf{\Phi}_{[1,1]} & \mathbf{0} \\ \mathbf{\Phi}_{[2,1]} & \mathbf{\Phi}_{[2,2]} \end{pmatrix} \\
=& \begin{pmatrix} \mathbf{S}_{t,[1,1]} + \left(\mathbf{I} - \mathbf{S}_{t,[1,1]}\right)\mathbf{\Phi}_{[1,1]} & \mathbf{0} \\ \left(\mathbf{I} - \mathbf{S}_{t,[2,2]}\right)\mathbf{\Phi}_{[2,1]} & \mathbf{S}_{t,[2,2]} + \left(\mathbf{I} - \mathbf{S}_{t,[2,2]}\right)\mathbf{\Phi}_{[2,2]} \end{pmatrix}.
\end{aligned}
$$

We can then write $\mathbf{X}^{\mathrm{Left}}_{t,[2,1]}$ recursively and substitute for $\mathbf{X}_t$ to yield

$$
\begin{aligned}
\mathbf{X}^{\mathrm{Left}}_{t,[2,1]} &= \left[\mathbf{X}_t\mathbf{X}^{\mathrm{Left}}_{t-1}\right]_{[2,1]} \\
&= \mathbf{X}_{t,[2,1]}\mathbf{X}^{\mathrm{Left}}_{t-1,[1,1]} && + \mathbf{X}_{t,[2,2]}\mathbf{X}^{\mathrm{Left}}_{t-1,[2,1]} \\
&= \left(\mathbf{I} - \mathbf{S}_{t,[2,2]}\right)\mathbf{\Phi}_{[2,1]}\mathbf{X}^{\mathrm{Left}}_{t-1,[1,1]} && + \left[\mathbf{S}_{t,[2,2]} + \left(\mathbf{I} - \mathbf{S}_{t,[2,2]}\right)\mathbf{\Phi}_{[2,2]}\right]\mathbf{X}^{\mathrm{Left}}_{t-1,[2,1]}.
\end{aligned}
$$

To simplify notation define

$$
\begin{aligned}
\mathbf{K} &\equiv \left[\mathbf{I} - \mathbf{\Phi}_{[2,2]}\right]^{-1}\mathbf{\Phi}_{[2,1]} \\
\mathbf{R}_t &\equiv \mathbf{S}_{t,[2,2]} + \left(\mathbf{I} - \mathbf{S}_{t,[2,2]}\right)\mathbf{\Phi}_{[2,2]},
\end{aligned}
$$

which implies

$$
\mathbf{I} - \mathbf{S}_{t,[2,2]} = (\mathbf{I} - \mathbf{R}_t)\left[\mathbf{I} - \mathbf{\Phi}_{[2,2]}\right]^{-1}.
$$

Substituting in the recursive equation above yields

$$\mathbf{X}^{\text{Left}}_{t,[2,1]} = \mathbf{R}_t \mathbf{X}^{\text{Left}}_{t-1,[2,1]} + (\mathbf{I} - \mathbf{R}_t)\, \mathbf{K}\mathbf{X}^{\text{Left}}_{t-1,[1,1]}.$$

Iterating this recursive definition of $\mathbf{X}^{\text{Left}}_{t,[2,1]}$ over t and simplifying by cancelling terms yields

$$\begin{aligned}
\mathbf{X}^{\text{Left}}_{t,[2,1]} &= \mathbf{R}_t \mathbf{X}^{\text{Left}}_{t-1,[2,1]} + (\mathbf{I} - \mathbf{R}_t)\, \mathbf{K}\mathbf{X}^{\text{Left}}_{t-1,[1,1]} \\
&= \mathbf{R}_t \left\{ \mathbf{R}_{t-1} \mathbf{X}^{\text{Left}}_{t-2,[2,1]} + (\mathbf{I} - \mathbf{R}_{t-1})\, \mathbf{K}\mathbf{X}^{\text{Left}}_{t-2,[1,1]} \right\} \\
&\quad + (\mathbf{I} - \mathbf{R}_t)\, \mathbf{K}\mathbf{X}^{\text{Left}}_{t-1,[1,1]} \\
&= \mathbf{R}_t \mathbf{R}_{t-1} \dots \mathbf{R}_1 \mathbf{X}^{\text{Left}}_{0,[2,1]} \\
&\quad + (\mathbf{I} - \mathbf{R}_t)\, \mathbf{K}\mathbf{X}^{\text{Left}}_{t-1,[1,1]} \\
&\quad + \mathbf{R}_t\, (\mathbf{I} - \mathbf{R}_{t-1})\, \mathbf{K}\mathbf{X}^{\text{Left}}_{t-2,[1,1]} \\
&\quad + \mathbf{R}_t \mathbf{R}_{t-1}\, (\mathbf{I} - \mathbf{R}_{t-2})\, \mathbf{K}\mathbf{X}^{\text{Left}}_{t-3,[1,1]} \\
&\quad + \\
&\quad \vdots \\
&\quad + \mathbf{R}_t \mathbf{R}_{t-1} \dots \mathbf{R}_2\, (\mathbf{I} - \mathbf{R}_1)\, \mathbf{K}\mathbf{X}^{\text{Left}}_{0,[1,1]} \\
&= \mathbf{K}\mathbf{X}^{\text{Left}}_{t-1,[1,1]} \\
&\quad + \mathbf{R}_t \mathbf{R}_{t-1} \dots \mathbf{R}_1 \left[\mathbf{X}^{\text{Left}}_{0,[2,1]} - \mathbf{X}^{\text{Left}}_{0,[1,1]} \right] \\
&\quad + \mathbf{R}_t \mathbf{K} \Delta \mathbf{X}^{\text{Left}}_{t-1,[1,1]} \\
&\quad + \mathbf{R}_t \mathbf{R}_{t-1} \mathbf{K} \Delta \mathbf{X}^{\text{Left}}_{t-2,[1,1]} \\
&\quad + \\
&\quad \vdots \\
&\quad + \mathbf{R}_t \mathbf{R}_{t-1} \dots \mathbf{R}_2 \mathbf{K} \Delta \mathbf{X}^{\text{Left}}_{1,[1,1]} \\
&= \mathbf{K}\mathbf{X}^{\text{Left}}_{t-1,[1,1]} \\
&\quad + \mathbf{R}_t \mathbf{R}_{t-1} \dots \mathbf{R}_1 \left[\mathbf{X}^{\text{Left}}_{0,[2,1]} - \mathbf{X}^{\text{Left}}_{0,[1,1]} \right]. \\
&\quad + \sum_{i=1}^{t-1} \left\{ \mathbf{R}_t \dots \mathbf{R}_{t-i+1} \mathbf{K} \Delta \mathbf{X}^{\text{Left}}_{t-i,[1,1]} \right\}
\end{aligned} \tag{4}$$

where

$$\Delta \mathbf{X}^{\text{Left}}_{t,[1,1]} = \mathbf{X}^{\text{Left}}_{t,[1,1]} - \mathbf{X}^{\text{Left}}_{t-1,[1,1]}$$

We discuss convergence of all three terms in Equation (4).

The first term converges to $\mathbf{K}\overline{\mathbf{X}}_{[1,1]} = \left[\mathbf{I} - \mathbf{\Phi}_{[2,2]}\right]^{-1} \mathbf{\Phi}_{[2,1]}\overline{\mathbf{X}}_{[1,1]}$. This follows from the argument regarding the block $\mathbf{X}^{\text{Left}}_{t,[1,1]}$ above.

The second term converges to the zero matrix. The expression is proportional to $\mathbf{R}_t\mathbf{R}_{t-1}\ldots\mathbf{R}_1$, which is the left product of a sequence of matrices $\mathbf{R}_i$. By its definition matrix $\mathbf{R}_i$ is the weighted average of the identity matrix and $\mathbf{\Phi}_{[2,2]}$ with strictly positive weight on $\mathbf{\Phi}_{[2,2]}$ for all t as the diagonal elements of $\mathbf{S}_i$ are strictly positive due to P14 Assumption 1. This implies that $\mathbf{R}_i$ inherits from $\mathbf{\Phi}_{[2,2]}$ the property that it is substochastic with at least one row adding up to strictly less than one and all individuals inessential. It then follows that the second term in Equation (4) converges to the zero matrix analogous to the argument for $\mathbf{X}^{\text{Left}}_{t,[2,2]}$ above.

Finally, consider the third term involving the sum of terms including $\Delta\mathbf{X}^{\text{Left}}_{t,[1,1]}$. We show that each summand has an upper bound in a matrix norm that converges to zero exponentially with t. Each summand indexed by $i \in \{1, 2, \ldots, t-1\}$ is the left product of $\mathbf{R}_t \ldots \mathbf{R}_{t-i+1}$ and $\Delta\mathbf{X}^{\text{Left}}_{t-i,[1,1]}$. We consider these in turn.

a) $\mathbf{R}_t \ldots \mathbf{R}_{t-i+1}$ Each matrix $\mathbf{R}_i$ is substochastic and consists of inessential individuals and thus its left product converges to zero. Moreover, from all individuals being inessential it follows that there exists k such that every row of $\mathbf{R}_{i+k}\mathbf{R}_{i+k-1}\ldots\mathbf{R}_i$ adds up to strictly less than one for any non-negative i. This implies that there exists $\beta_R < 1$ such that $\|\mathbf{R}_{i+k}\mathbf{R}_{i+k-1}\ldots\mathbf{R}_i\| \leq \beta_R$ where $\|\cdot\|$ is the maximum row sum or ∞-norm defined by $\|A\| = \max_i \left\{\sum_j |a_{ij}|\right\}$. For any i we thus have $\|\mathbf{R}_t\mathbf{R}_{t-1}\ldots\mathbf{R}_{t-i+1}\| \leq \beta_R^{\lfloor i/k \rfloor} \leq \beta_R^{(i-k)/k} = \left(\beta_R^{1/k}\right)^{i-k}$ which is strictly less than one for any $i \geq k$.

b) $\Delta\mathbf{X}^{\text{Left}}_{t-i,[1,1]}$ For any given i, $\|\Delta\mathbf{X}^{\text{Left}}_{t-i,[1,1]}\|$ can be bounded from above using the triangle inequality and the second part of Theorem 1.10 Hartfiel (2006).

$$
\begin{aligned}
\|\Delta\mathbf{X}^{\text{Left}}_{t-i,[1,1]}\| &= \|\mathbf{X}^{\text{Left}}_{t-i,[1,1]} - \mathbf{X}^{\text{Left}}_{t-i-1,[1,1]}\| \\
&= \| \left[\mathbf{X}^{\text{Left}}_{t-i,[1,1]} - \overline{\mathbf{X}}_{[1,1]}\right] - \left[\mathbf{X}^{\text{Left}}_{t-i-1,[1,1]} - \overline{\mathbf{X}}_{[1,1]}\right] \| \\
&\leq \| \left[\mathbf{X}^{\text{Left}}_{t-i,[1,1]} - \overline{\mathbf{X}}_{[1,1]}\right] \| + \| \left[\mathbf{X}^{\text{Left}}_{t-i-1,[1,1]} - \overline{\mathbf{X}}_{[1,1]}\right] \| \\
&\leq K_X\beta_X^{t-i} + K_X\beta_X^{t-i-1} \\
&= K_X\beta_X^{t-i-1}(\beta_X + 1)
\end{aligned}
$$

for some constants K_X and $\beta_X < 1$ following Theorem 1.10 Hartfiel (2006).

It then follows that we can bound the ∞-norm of every summand from above as follows.

$$\begin{aligned}
\|\mathbf{R}_t \ldots \mathbf{R}_{t-i+1}\mathbf{K}\Delta\mathbf{X}^{\text{Left}}_{t-i,[1,1]}\| &\leq \|\mathbf{R}_t \ldots \mathbf{R}_{t-i+1}\|\,\|\mathbf{K}\|\,\|\Delta\mathbf{X}^{\text{Left}}_{t-i,[1,1]}\| \\
&\leq \left[\beta_R^{1/k}\right]^{i-k} \|\mathbf{K}\| K_X \beta_X^{t-i-1}(\beta_X + 1) \\
&\leq \beta^{i-k}\|\mathbf{K}\| K_X \beta^{t-i-1}(\beta+1) \\
&\leq \|\mathbf{K}\| K_R K_X \beta^{t-1-k}(\beta+1) \\
&= \lambda\beta^{t-1}
\end{aligned}$$

for $\beta = \max\{\beta_R^{1/k}, \beta_X\} < 1$ and some scalar λ that collects terms invariant with t. This bound applies to all summands in the sum in Equation (4). By the sub-additivity of the norm $\|\cdot\|$ and L'Hôpital's rule

$$\begin{aligned}
\lim_{t\to\infty} \left\| \sum_{i=1}^{t-1} \left\{\mathbf{R}_t \ldots \mathbf{R}_{t-i+1}\mathbf{K}\Delta\mathbf{X}^{\text{Left}}_{t-i,[1,1]}\right\} \right\| &\leq \lim_{t\to\infty} \sum_{i=1}^{t-1} \|\mathbf{R}_t \ldots \mathbf{R}_{t-i+1}\mathbf{K}\Delta\mathbf{X}^{\text{Left}}_{t-i,[1,1]}\| \\
&\leq \lim_{t\to\infty} \sum_{i=1}^{t-1} \lambda\beta^{t-1} \\
&= \lambda \lim_{t\to\infty}(t-1)\beta^{t-1} \\
&= 0 \text{ for } \beta < 1.
\end{aligned}$$

Therefore the infinite sum forming the third term converges to the zero matrix.

Summarizing, we have

$$\begin{aligned}
\lim_{t\to\infty} \mathbf{X}^{\text{Left}}_{t,[2,1]} &= \lim_{t\to\infty} \mathbf{K}\mathbf{X}^{\text{Left}}_{t-1,[1,1]} \\
&= \left[\mathbf{I} - \mathbf{\Phi}_{[2,2]}\right]^{-1} \mathbf{\Phi}_{[2,1]}\overline{\mathbf{X}}_{[1,1]}
\end{aligned}$$

as required. □

This new proof includes a characterization of the long-run traits for inessential individuals as a weighted average over the long-run traits of the essential individuals. Furthermore, those weights are *independent* of the sequence of vertical socialization weights $\mathbf{S}_{t,[2,2]}$ for the inessential individuals. Thus, the weights assigned to oblique socialization as captured in $\mathbf{\Phi}_{[2,1]}$ and $\mathbf{\Phi}_{[2,2]}$ are sufficient to describe the long-run traits of individuals in the inessential group relative to the traits of the essential individuals. The convergence result in P14 can thus be extended to include this characterization.

(P14) Proposition 2 (Extended) The system described by Equations (1) and (2) converges for any time-invariant row stochastic matrix $\mathbf{\Phi}$. The long-run traits of the inessential individuals converge to a weighted average of the long-run traits of the essential classes with weights given by $\left[\mathbf{I} - \mathbf{\Phi}_{[2,2]}\right]^{-1} \mathbf{\Phi}_{[2,1]}$. The long-run traits of inessential individuals are independent of the time-varying matrix of socialization efforts $\mathbf{S}_t$.

4 Discussion

There are two main points that we aim to highlight in this note. First, for time-varying transition matrices the convergence behaviour of repeated averaging models is not identical to that of Markov chains. Section 1 highlight this point by providing a stark example of a sequence of row stochastic matrices that converges when multiplied from the right as a Markov chain but that enters a limit cycle when multiplied from the left as in a cultural traits model. The difference between the two settings is also reflected in an established literature on the mathematics of repeated averaging models that – while acknowledging parallels with Markov chains – offers independent convergence results for left multiplications.

It is worth noting that time variation is a necessary condition for this issue to arise. If the sequence of transition matrices is not time-varying then results from Markov chain theory readily translate into the opinion dynamics context. For example, Theorem 8.1 in Matthew O Jackson (2008) and results in Golub and Matthew O. Jackson (2010) provide conditions for convergence with invariant transition matrices that draw directly on Markov chain theory. In the example in Section 1, if the sequence of matrices is altered to a sequence of $\mathbf{X}_{\text{odd}}$ or $\mathbf{X}_{\text{even}}$ only, so that it is no longer time-varying, then there is convergence both with multiplication from the left and the right. This may also be the reason why the distinction between Markov chains and cultural traits models appears insufficiently appreciated in the literature.

The second more general message is that convergence in cultural traits transmission or opinion dynamics models with time-varying transition matrices generally requires relatively strong assumptions to prevent cycling. For example, in the case of P14 Proposition 2, convergence is ensured by the structure of the transition matrix embodied in Equation (2) which restricts time variation to the weight each individual assigns to her own traits. This property plays a critical part of the new proof included in this note. The P14 model is itself a generalization of the setting in DeMarzo, Vayanos, and Zwiebel (2003) who study a setting with time variation in the weight on own beliefs that is restricted to be the same across all individuals in any period. Corollary 1 exploits a different set of conditions, specifically those of Theorem 1.10 Hartfiel (2006), that there is exactly one essential class with a regular transition matrix. By contrast, Büchel, Hellmann, and Pichler (2014) pursue a different approach to ensure convergence for the general case with time-varying transition matrices they study in Appendix C. They impose a certain form of symmetry on the socialization matrix and then build on results for the convergence of left products of matrices in Lorenz (2005) and Lorenz (2006). Finally, Prummer and Siedlarek (2017) present a model of cultural transmission with community leaders, that are in effect two isolated essential individuals, and a group of followers who assign time-varying weight on the traits of the leaders. They show that convergence is ensured under their Assumption 1 that limits the speed with which weights change from one period to the next. Their Proposition 2.2 establishes that under this assumption the cultural traits updating process is a contraction and thus converges globally to a unique steady state.

References

Büchel, Berno, Tim Hellmann, and Michael M Pichler (2014). "The Dynamics of Continuous Cultural traits in Social Networks". *Journal of Economic Theory* 154, pp. 274–309. DOI: 10.1016/j.jet.2014.09.008.

Cavalli-Sforza, Luigi L and Marcus W Feldman (1973). "Cultural versus Biological Inheritance: Phenotypic Transmission from Parents to Children.(A Theory of the Effect of Parental Phenotypes on Children's Phenotypes)." *American Journal of Human Genetics* 25.6, p. 618.

D'Amico, Guglielmo, Jacques Janssen, and Raimondo Manca (2009). "The Dynamic Behaviour of Non-Homogeneous Single-unireducible Markov and Semi-Markov Chains". In: *Networks, Topology and Dynamics*. Springer, pp. 195–211. DOI: 10.1007/978-3-540-68409-1_10.

DeGroot, Morris H (1974). "Reaching a Consensus". *Journal of the American Statistical Association* 69.345, pp. 118–121. DOI: 10.1080/01621459.1974.10480137.

DeMarzo, Peter M, Dimitri Vayanos, and Jeffrey Zwiebel (2003). "Persuasion Bias, Social Influence, and Unidimensional Opinions". *The Quarterly Journal of Economics* 118.3, pp. 909–968. DOI: 10.1162/00335530360698469.

Golub, Ben and Matthew O. Jackson (2010). "Naive Learning in Social Networks and the Wisdom of Crowds". *American Economic Journal: Microeconomics* 2.1, pp. 112–149. DOI: 10.1257/mic.2.1.112.

Hartfiel, Darald J (2006). *Markov Set-Chains*. Springer. DOI: 10.1007/bfb0094586.

Jackson, Matthew O (2008). *Social and Economic Networks*. Princeton University Press.

Lorenz, Jan (2005). "A Stabilization Theorem for Dynamics of Continuous Opinions". *Physica A: Statistical Mechanics and its Applications* 355.1, pp. 217–223. DOI: 10.1016/j.physa.2005.02.086.

Lorenz, Jan (2006). "Convergence of Products of Stochastic Matrices with Positive Diagonals and the Opinion Dynamics Background". In: *Lecture Notes in Control and Information Sciences*. Vol. 341. Springer, pp. 209–216.

Panebianco, Fabrizio (2014). "Socialization Networks and the Transmission of Interethnic Attitudes". *Journal of Economic Theory* 150, pp. 583–610. DOI: 10.1016/j.jet.2013.09.003.

Prummer, Anja and Jan-Peter Siedlarek (2017). "Community Leaders and the Preservation of Cultural Traits". *Journal of Economic Theory* 168, pp. 143 –176. ISSN: 0022-0531. DOI: 10.1016/j.jet.2016.12.007.