Research Article

Dynamics from Multivariable Longitudinal Data

Maria Vivien Visaya and David Sherwell

School of Computational and Applied Mathematics, University of the Witwatersrand (Wits), Private Bag 3, Johannesburg 2050, South Africa

Correspondence should be addressed to Maria Vivien Visaya; v3isaya@gmail.com

Received 15 July 2013; Revised 12 December 2013; Accepted 15 December 2013; Published 19 March 2014

Academic Editor: Mitsuhiro Ohta

Copyright © 2014 M. V. Visaya and D. Sherwell. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We introduce a method of analysing longitudinal data in $n \geq 1$ variables and a population of $K \geq 1$ observations. Longitudinal data of each observation is exactly coded to an orbit in a two-dimensional state space S_n . At each time, information of each observation is coded to a point $(x, y) \in S_n$, where x is the physical condition of the observation and y is an ordering of variables. Orbit of each observation in S_n is described by a map that dynamically rearranges order of variables at each time step, eventually placing the most stable, least frequently changing variable to the left and the most frequently changing variable to the right. By this operation, we are able to extract dynamics from data and visualise the orbit of each observation. In addition, clustering of data in the stable variables is revealed. All possible paths that any observation can take in S_n are given by a subshift of finite type (SFT). We discuss mathematical properties of the transition matrix associated to this SFT. Dynamics of the population is a nonautonomous multivalued map equivalent to a nonstationary SFT. We illustrate the method using a longitudinal data of a population of households from Agincourt, South Africa.

1. Introduction

Analysis of multivariable longitudinal data involves either statistical or nonstatistical methods. Statistical methods include multivariate Markov chain model [1], regression model [2], and mixed models [3], while some nonstatistical methods involve extraction of dynamical system using state space reconstruction technique [4] or visual methods such as motion charts [5, 6] and parallel coordinate plots [7, 8]. A motion chart shows additional dimensions of the data at different time points, where the size and color of the bubble (among others) are used as variables. PCP represents variables as parallel axes, where a sequence of line segments intersects each axis at a point corresponding to the observation's value at the associated variable. Both methods aim to identify correlation among variables and identification of clusters and patterns among observations in the data. We present a novel nonstatistical method of analysis that is useful particularly for collecting information of change. State space reconstruction method mainly requires use of delay coordinates from data to come up with models for prediction. Our method does not rely on delay coordinates, and our aim is not prediction. Contrasting to motion charts, there is, in

principle, no limitation to the number of variables studied in our method. Contrasting to PCP, for large n and large number of observations, orbits over our two-dimensional state space, or over time, are easily visualized.

Here we consider dynamics of multivariable longitudinal data of a population of observations by applying a swap operation on the data of each observation. We suppose that values of $n \geq 1$ variables are discretised to $m > 1$ bins so that each variable takes value from $V = \{0, 1, \dots, m-1\}$. For a fixed order of variables (e.g., $v_0, v_1, \dots, v_{n-1}$) denote by

$$M_{n,m} = V^n \quad (1)$$

the space of all m -ary sequences of length n and element $x \in M$ by $x = x_0 x_1 \dots x_{n-1}$ such that x_i is the value of v_i . The m -ary multivariable time series in $n \geq 1$ variables of an observation k is defined by

$$X^k = \{X_0^k, X_1^k, \dots, X_T^k\}, \quad (2)$$

where each $X_t^k \in M_{n,m}$. Our method can be used for m -ary valued data, but for illustration, we use $m = 2$. We denote our binary multivariable longitudinal data D of a population of K

TABLE 1: Unfavourable = 0/favourable = 1 coding of four households to questions related to biological mother (BM), household head (HH), and death (AD).

Time	$h1$ BM, HH, AD	$h2$ BM, HH, AD	$h3$ BM, HH, AD	$h4$ BM, HH, AD
0	0, 1, 1	0, 1, 1	0, 0, 1	0, 1, 1
1	1, 1, 0	0, 0, 1	0, 0, 1	1, 0, 1
2	0, 1, 1	1, 1, 1	1, 0, 1	1, 1, 1
3	0, 1, 1	1, 0, 0	0, 0, 0	0, 0, 1
4	0, 1, 1	0, 1, 0	0, 0, 1	0, 1, 1
5	0, 1, 1	0, 1, 0	0, 0, 1	1, 0, 1
6	1, 1, 1	0, 0, 0	1, 0, 1	0, 0, 1

observations by the set of binary multivariable time series of observations

$$D = \{X^1, X^2, \dots, X^K\}. \quad (3)$$

In the general analysis of longitudinal data we see no recognition that data is taken with a purpose. Here we suppose that the value of a binary variable is either favourable or unfavourable to a given purpose. Suppose we would like to investigate the effect of change of households variables, namely, biological mother (BM), household head (HH), and adult death (AD), to a child's educational progress. Consider $n = 3$ binary questions

- Q_0 : Is the biological mother present? (BM)
 Q_1 : Is the household headed by a minor? (HH) (4)
 Q_2 : Is there an adult death in the household? (AD)

and suppose that answers are coded either favourable = 1 or unfavourable = 0 to our purpose. It is a reasonable hypothesis that the answer “yes” to Q_0 is favourable to child education, while “yes” to Q_1 and Q_2 is unfavourable. Table 1 shows coded data of 4 subjects (households) over 7 time steps. Using parallel coordinate plots (PCP), Figure 1(a) shows the answer of each subject at $t = 0$. To illustrate the answer for all time t using PCP, the evolution of line segments, with time, will form surfaces that obscure each other. Figure 1(b) shows the Bratelli diagram [9, 10] of transition of answers from time t to $t + 1$, to variable order BM, HH, and AD. Transitions per time interval can be associated with a state transition matrix, where states are analysed on the space

$$M_3 = \{000, 001, 010, 011, 100, 101, 110, 111\}. \quad (5)$$

This can be used to generate probability matrices for Markov models [11, 12].

Observe that, for question regarding Q_1 (HH), household $h1$ has constant favourable answer, while $h3$ has constant unfavourable answer. On the other hand, $h4$ has constant favourable answer to Q_3 (AD). Our aim is to extract clusters associated with stable variables. Underlying our method is our belief that the set of physical variables in which an observation spends the most time in is important (e.g., HH

= 1 for $h1$, HH = 0 for $h3$, and AD = 1 for $h4$) and that among the physical variables themselves the variables that are most often experienced (most probable) by the observations are important. We elegantly expose both most probable variable and value of variable, by a simple process of dynamically reordering variables.

There is no a priori indication of any absolute dynamics in data and here it is deterministically imposed. Because longitudinal data is fundamentally defined by change (if nothing changes, cross-sectional data is sufficient), frequency of answer change of variables then becomes a property of interest. A deterministic operation is applied to the multivariable data of each observation at each time step, dynamically reordering position of variables (and their corresponding values) by their stability; that is, the most stable is eventually positioned to the left, and the most frequently changing one is positioned to the right. All possible orders of variables are considered. From this, we introduce the significance state (order of variables) and fitness state (associated values with variables) of an observation. It is in a chosen ordering that the notion of fitness takes an objective and consistent meaning. The idea of fitness and significance is new in literature. The n -dimensional longitudinal data of an observation is represented as a 2-dimensional orbit in fitness-significance axes S_n composed of $2^n n!$ points. Orbits in S_n sufficiently encode the longitudinal data of each observation. Analysis of orbits at the individual and population levels in this space can then follow.

This paper is organized as follows. In Section 2, we present the method of constructing orbits from multivariable longitudinal data. A detailed theory of the reordering operation applied to data is presented. In Section 3, a deterministic equation of motion that generates all observed orbits, as well as other possible orbits, is presented. In Section 4 we discuss transitions that occur in data. This is captured by a nonstationary SFT. We also discuss dynamics at the population level. An illustration of the method is presented in Section 5. We give concluding remarks in Section 6.

2. Background and Preliminaries: Method of Orbits

We suppose that longitudinal data is gathered by first specifying a purpose and then choosing questions that are of interest to purpose. The questionnaire may be designed to a purpose posed in the form “to investigate the effect of the n variables $v_0, v_2, \dots, v_{n-1}$ on Z .” Here we will only consider binary-valued questions with responses hypothesized as either favourable or unfavourable to the purpose. A favourable answer is coded 1 and 0 otherwise. Our longitudinal data is the response to a set of $n \geq 1$ questions (associated with variables) surveyed from a population P composed of $K \geq 1$ observations over T periods.

Denote the *questionnaire* by

$$Q = \{Q_0, Q_1, \dots, Q_{n-1}\}, \quad (6)$$

the set containing n questions. Let $I_n = \{0, 1, \dots, n - 1\}$ be an index set of n elements, and let time $t = 0, 1, \dots, T$.

FIGURE 1: Visual display of data of four subjects in Table 1 (a) at $t = 0$ using PCP, where each parallel line is composed of answers 0 and 1 and (b) for all time $t = 0, 1, \dots, 6$ using a Bratelli diagram, where each parallel line is composed of states from $M_3 = \{0, 1\}^3$, the space of binary sequences of length 3.

TABLE 2: Data of observation k (column 2), with questions reordered at $t = 0$ and $t = 1$.

Observation k	Coded answers to $Q = \{Q_0, Q_1, Q_2\}$	Concatenation of answers to Q_t^k	Concatenation of indices of Q_t^k
$t = 0$	$\{0, 1, 0\}$	100	102
1	$\{0, 0, 1\}$	010	021

Let $j, i_j \in I_n$ and let $x_j \in \{0, 1\}$. For each observation $k = 1, 2, \dots, K$, denote a reordering of questions in Q at time t by

$$Q_t^k = \{Q_{i_0 t}^k, Q_{i_1 t}^k, \dots, Q_{i_{(n-1)t}^k}^k\} \quad (7)$$

$$:= \{Q_{i_0}, Q_{i_1}, \dots, Q_{i_{n-1}}\}_t^k,$$

the concatenation of answers to Q_t^k by

$$x_t^k = x_{0t}^k x_{1t}^k \cdots x_{(n-1)t}^k = (x_{jt}^k)_{j=0}^{n-1} \quad (8)$$

$$:= (x_0 x_1 \cdots x_{(n-1)})_t^k,$$

and the concatenation of question indices of Q_t^k by

$$y_t^k = i_{0t}^k i_{1t}^k \cdots i_{(n-1)t}^k = (i_{jt}^k)_{j=0}^{n-1} \quad (9)$$

$$:= (i_0 i_1 \cdots i_{(n-1)})_t^k.$$

Example 1. Coded data of observation k for $t = 0$ and $t = 1$ is shown in column 2 of Table 2. Suppose we (arbitrarily) reorder the questions at $t = 0$ to $\{Q_1, Q_0, Q_2\}$. Then $x_0^k = 010$ and $y_0^k = 102$. Similarly, if $Q_1^k = \{Q_2, Q_0, Q_1\}$, we have $x_1^k = 001$ and $y_1^k = 201$. As we are merely rewriting entries from the original data, all information is preserved.

2.1. Fitness and Significance States. Consider questions in (4) and assign index i to Q_i ($i = 0, 1, 2$). Suppose we give more weight (significance) to Q_0 and do this by positioning 0 in the left-end of question order, say, Q_0, Q_1, Q_2 , denoted by 012 (or $Q_0, Q_2, Q_1 = 021$). As in numbers or decimals, our weighting places the most significant number at the left-end position. All possible (concatenated) answers to 012 are given by

$$001, 000, 010, 111, 110, 101, 100, 011. \quad (10)$$

TABLE 3: Concatenated coded answers of observations to question order $Q_0, Q_1, Q_2 = (012)$.

t	k	k'	k''
0	010	100	111
1	001	110	001
2	000	101	001
3	010	110	111
4	000	100	001

Since 000 has all unfavourable answers we say that it is the least *fit* answer, while 111 is the *fittest*. Note that there are states with the same number of favourable values, for example, 001, 100, and 010. By a suitable weighting of questions, we show below that the lexicographic ordering of answers in (10) is an appropriate ordering of fitness.

Table 3 illustrates concatenated coded answers to question order 012 of three observations from a population. To Q_0 , observation k has constant answer 0 while k' has constant answer 1.

Suppose we arrange answers in (10) lexicographically along an x -axis. Then for question order 012, a one-dimensional dynamics on the x -axis composed of the eight states arises. Answers of k and k' to Q_0 seldom change (i.e., they are both constant in Q_0) so the two households k and k' stay in the regions $0**$ and $1**$ of the x -axis, respectively. Recall that Q_0 is the question associated with the significant (left) position of the question order so fitness is biased towards the left position. We can then write $x = 0** < x' = 1**$ because the significant variable Q_0 is unfavourable in x and favourable in x' . This holds true even if x has the same, or more, favourable values as x' . This argument can be extended to any two elements $x, x' \in M_n$ with the same first ℓ entries.

In general, not all observations may be stable in the same variable; for example, k'' in Table 3 is constant in Q_0 , not

in Q_2 . Moreover, stability of an observation may change in time; it may be stable in variable i over one time interval and then stable in variable i' over another time interval. We will not study orbits in a fixed question order alone. We construct a y -axis with states corresponding to question orders. The order of questions per observation becomes a new variable.

Definition 2. Given $n \geq 1$ and Q_k^t , the *fitness state* and *significance state* of observation k at time t are the sequences

$$x_t^k = (x_{jt})_{j=0}^{n-1}, \quad y_t^k = (i_{jt})_{j=0}^{n-1}, \quad (11)$$

respectively. The set of fitness states of length n is called the *fitness space* of n variables defined by

$$X_n := \{(x_j)_{j=0}^{n-1} : x_j \in \{0, 1\}\} \quad (12)$$

and the set of significance states of length n is called the *significance space* of n variables defined by

$$Y_n := \{(i_j)_{j=0}^{n-1} : i_j \in I_n, i_j \text{'s distinct}\}. \quad (13)$$

Elements of both X_n and Y_n are arranged according to the lexicographic ordering ($<$) of sequences of length n .

Definition 3. The space

$$\begin{aligned} S_n &= \{p = (x, y) : x \in X_n, y \in Y_n\} \\ &= X_n \times Y_n \end{aligned} \quad (14)$$

is called the *change space* for n variables.

Given n , we have the cardinalities $|X_n| = 2^n$, $|Y_n| = n!$, and $|S_n| = N = 2^n n!$. A way of labeling state $s_j \in S_n$ is via the map

$$\begin{aligned} \psi : S_n &\longrightarrow \{1, 2, \dots, N\}, \quad s_j \longmapsto \psi(s_j) = j, \\ j &= 1, 2, \dots, N. \end{aligned} \quad (15)$$

For convenience, we label states in S_n from left to right, and from top to bottom. If $s_j = (x, y)$ and $\psi(s_j) = j$, we will refer to state s_j as state j and write $j = (x, y)$.

Remark 4. In general, for multivariate data in n variables, with *all* variables m -ary valued, the space $S_{m,n}$ is composed of $m^n \times n!$ states.

2.2. The Method of Orbits. We define the dynamics of observations taken from a survey of $n \geq 1$ questions. Let $\mathbb{N}_0$ be the set of nonnegative integers, let $\mathcal{C}_n = \mathcal{P}(I_n)$ be the power set of I_n , and let x_j^* be such that

$$x_j^* = \begin{cases} 1 & \text{if } x_j = 0 \\ 0 & \text{if } x_j = 1. \end{cases} \quad (16)$$

Definition 5. Let $\Delta \in \mathcal{C}_n$ and let $j \in \Delta$.

(a) The map $\phi_\Delta : S_n \rightarrow S_n$ is defined by

$$\phi_\Delta(x_0 x_1 \cdots x_j \cdots x_{n-1}, y) = (x_0 x_1 \cdots x_j^* \cdots x_{n-1}, y). \quad (17)$$

(b) The map $\phi_J : S_n \rightarrow S_n$ is defined by

$$\begin{aligned} \phi_J(x_0 x_1 \cdots x_j \cdots x_{n-1}, i_0 i_1 \cdots i_j \cdots i_{n-1}) \\ = (x_0 x_1 \cdots x_{j-1} x_{j+1} \cdots x_{n-1} x_j, i_0 i_1 \cdots i_{j-1} i_{j+1} \cdots i_{n-1} i_j). \end{aligned} \quad (18)$$

Let $j, j' \in \Delta$. Then x_j and $x_{j'}$ both change values under ϕ_Δ . If $j < j'$, then ϕ_J is first applied to j' ; that is,

$$\begin{aligned} \phi_J(x_0 x_1 \cdots x_j \cdots x_{j'} \cdots x_{n-1}, i_0 i_1 \cdots i_j \cdots i_{j'} \cdots i_{n-1}) \\ = (x_0 x_1 \cdots x_{j-1} x_{j+1} \cdots x_{j'-1} x_{j'+1} \cdots x_{n-1} x_{j'} x_j, \\ i_0 i_1 \cdots i_{j-1} i_{j+1} \cdots i_{j'-1} i_{j'+1} \cdots i_{n-1} i_{j'} i_j). \end{aligned} \quad (19)$$

Definition 6. Let $i_j \in I_n$. Consider questionnaire Q with $Q_{i_j} \in Q$. For each observation k , let $f_{i_j}^k$ be the frequency of change in answer value of Q_{i_j} over the observation period. Suppose

$$0 < f_{i_0}^k < f_{i_1}^k < \cdots < f_{i_j}^k < \cdots < f_{i_{n-1}}^k. \quad (20)$$

Inequality (20) is called the *observation frequency relation* and question order $y_0^k = i_0 i_1 \cdots i_{n-1}$ is the *initial significance state* of observation k . If $f_{i_j}^k = f_{i_{j+1}}^k$ and $f_{i_j}^P < f_{i_{j+1}}^P$ at the population level, then choose question order $i_j i_{j+1}$. If $f_{i_j}^P = f_{i_{j+1}}^P$, then choose question order as in the questionnaire (6). The *initial fitness state* $x_0^k = x_0 x_1 \cdots x_{n-1}$ is such that x_j is the value of i_j in y_0^k . The *initial state* of k is the ordered pair $p_0^k = (x_0^k, y_0^k)$.

By choosing the initial significance y_0^k given in Definition 6, we start the orbit in its most-likely significance state. This facilitates convergence to clusters (where they exist) and is useful for short data sets. Other strategies (e.g., using order of (6) or random choice) will nonetheless converge to question order according to (20).

Remark 7. Longitudinal data is only of interest where change occurs; else cross-sectional surveys are adequate. We are interested in longitudinal data that give nontrivial information of change about the population; that is, $f_{i_j}^P \neq 0$. Otherwise, question i_j may be deleted as any such property becomes an identifier of subpopulations of possible interest for analysis in its own right.

Definition 8. For each observation k , define the *change set* at time t by

$$\begin{aligned} \Delta_t^k(\in \mathcal{C}_n) \\ = \{j : Q_{i_j} \in Q_t^k \text{ changes answer value at time } t+1\}. \end{aligned} \quad (21)$$

Let $p_0^k = (y_0^k, x_0^k)$ be the initial state of k . Denote by $p_t^k = (x_t^k, y_t^k)$ the state of k at time $t \geq 0$. The *change map* $\varphi : (\mathbb{N}_0, \mathcal{C}_n, S_n) \rightarrow S_n$ is such that

$$\begin{aligned} \varphi(t, \Delta_t^k, p_t^k) &:= \varphi_{[\Delta_t^k]}(p_t^k) \\ &= (\phi_J \circ \phi_{\Delta_t^k})(p_t^k) \\ &= p_{t+1}^k. \end{aligned} \quad (22)$$

The set $\Delta_t^k \in \mathcal{C}_n$ is given by the longitudinal data for each $t \geq 0$ and is a useful ordered listing of questions that change answer values from time t to $t + 1$. For each k , the nonautonomous map $\varphi_{[\Delta_t^k]}$ defines an evolutionary process that displaces the most frequently (resp., slowly) changing answers and corresponding questions to the right (resp., left).

Definition 9. Given initial state p_0^k of k , define the state of k at time $t \geq 1$ by

$$p_t = \varphi_{[\Delta_{t-1}]}(\dots(\varphi_{[\Delta_0]}(p_0))) . \quad (23)$$

The *forward orbit* of k under φ is defined by

$$\mathcal{O}_\varphi(k) = \{p_t^k\}_{t \in \mathbb{N}_0}. \quad (24)$$

2.3. Algorithm for Building the Orbit of an Observation in S_n . We give a simple algorithm to determine the states $p_{t \geq 0}^k$ that comprise the orbit of an observation k from longitudinal data.

Step 1. Determine initial question order y_0^k and initial state $p_0^k = (x_0^k, y_0^k)$ of k and plot p_0^k in S_n .

Step 2. Identify from the data of k the question that changes answer at time $t = 1$, say Q_{i_j} . Swap both i_j (in y_0^k) and corresponding answer x_j (in x_0^k) to the right of y_0^k and x_0^k , respectively, and change x_j to x_j^* , where $x_j^* = 0$ if $x_j = 1$ and 0 otherwise. This new question order and answer order give the next state $p_1^k = (x_1^k, y_1^k)$ of k . Suppose both Q_{i_j} and $Q_{i_{j'}}$ change answers at $t = 1$. If $j < j'$, then sequentially swap to the right i_j and $i_{j'}$ (resp., x_j and $x_{j'}$) of the question order (resp., answer order), starting with $i_{j'}$ (resp., $x_{j'}$). Change x_j to x_j^* and $x_{j'}$ to $x_{j'}^*$. Plot the point p_1^k in S_n and directed edge from p_0^k to p_1^k .

Step 3. Repeat Step 2, updating state $p_{t-1}^k \rightarrow p_t^k$ and $t \rightarrow t + 1$ for time $t = 1, 2, \dots, T - 1$.

Visualization and analysis of orbits of observations in S_n allow capturing information of change in longitudinal data. We illustrate in Figure 2 an orbit in S_3 . The useful distance on S_n is given by the discrete metric; that is, $d(p, p') = 1$ if $p \neq p'$ and zero otherwise. The visualized distance between points in S_n has no interpretation so we may represent S_n by a regularly spaced point.

FIGURE 2: State transitions of observation k (from Table 4) in S_3 .

TABLE 4: States of observation k .

Unit k	Coded answers to $Q = \{Q_0, Q_1, Q_2\}$	Fitness x_t^k	Significance y_t^k	Δ_t^k
$t = 0$	$\{1^*, 0^*, 0\}$	010	201	$\{1, 2\}$
1	$\{0, 1^*, 0\}$	010	210	$\{1\}$
2	$\{0^*, 0^*, 0^*\}$	000	201	$\{0, 1, 2\}$
3	$\{1, 1, 1\}$	111	102	

Remark 10. The set of all orderings of variables is captured in S_n . For fixed question order and $x, x' \in X_n$, $x < x'$ means that fitness state x' is fitter than x . Each level in the significance axis is question order under frequency ordering. The significance axis informs us which variables are weighted most strongly at each time, where significance is ordered from left to right. Clearly, reordering of variables is one among many families of operations; for example, swapping can be done by swapping changing variable to the left end. This operation however does not reveal clusters.

2.4. Transitions in S_n . We now analyze possible state transitions which an observation k can take in S_n .

Definition 11. Let $p = (x, y)$ and $p' = (x', y')$ be in S_n , and let $\Delta \in \mathcal{C}_n$.

- (a) Suppose Δ is such that $\varphi_{[\Delta]}(p) = p'$. Then there is a *transition* from p to p' under Δ , defined by $p \xrightarrow{[\Delta]} p'$.
- (b) Suppose $p \xrightarrow{[\Delta]} p'$. If $y = y'$ and $x \neq x'$, then there is a *horizontal transition* from p to p' . If $x = x'$ and $y \neq y'$, then there is a *vertical transition* from p to p' .
- (c) Suppose $p \xrightarrow{[\Delta]} p'$. If, in addition, Δ is such that $p' \xrightarrow{\Delta} p$, then the transition from p to p' is *reversible*, and one writes $p \xleftrightarrow{[\Delta]} p'$.

We have *self-transitions* if $\Delta = \emptyset$ (no change), the empty set. In general, horizontal transitions denote change in

FIGURE 3: (a) Orbits of the four subjects in Table 1. A red edge denotes a transition to the left, a green edge is a transition to the right, and a blue edge is a self-transition.

the right-most variable, while vertical transitions denote change in the last two variables.

Let $p, p' \in S_n$. Given an observation k , if $p_t^k \in p$ and $p_{t+1}^k \in p'$, then there is $\Delta_t \in \mathcal{C}_n$ such that $p \xrightarrow{[\Delta_t]} p'$. We use the symbol “ ϵ ” to denote that there may be other observations in p or p' at times t and $t + 1$, respectively. State transitions of observation k in S_n are visualised as a sequence of directed edges.

Example 12. Consider data of observation k in column 2 of Table 4. The asterisks denote changing answers in the next time step. The frequencies of change are $f_0^k = 2$, $f_1^k = 3$, and $f_2^k = 1$, so the initial question order of k is $y_0^k = 201$. And initial fitness state is $x_0^k = 010$ (Definition 6). At $t = 0$, we apply $\varphi_{[\Delta_0^k]}$ to initial state $p_0^k = (010, 201)$. Questions $i_1 = 0$ and $i_2 = 1$ in Q_0^k change answers in the next time step so $\Delta_0^k = \{1, 2\}$ (not $\{0, 1\}$). Applying $\phi_{\Delta_0^k}$ to $(010, 201)$ changes $x_2 = 0$ and $x_1 = 1$ to $x_2^* = 1$ and $x_1^* = 0$, respectively. Next, ϕ_j is first applied to both x_2^* and i_2 by moving each to the right (they are already rightmost), followed by moving x_1^* and i_1 to the right end. Hence, we have $(x, y)_1^k = (010, 210)$. For $t = 1$ and $t = 2$, verify that $\Delta_1^k = \{1\}$ and $\Delta_2^k = \{0, 1, 2\}$, respectively. At each time t the bold numbers in the significance column are the question indices in Q that change answers at $t + 1$. The orbit of k in S_3 is shown in Figure 2. The vertical transition from state 11 to 3 denotes two changing answers, while a transition from state 9 to 32 denotes three changing answers.

Example 13. Figures 3(a) and 3(b) illustrate the associated orbits in S_3 of the four households given in Table 1. Observe that orbit of $h3$ stays strictly on the left half of S_3 , particularly in subset where $Q_1(\text{HH}) = 0$, while $h1$ and $h4$ stay in the right half subset, where $Q_1(\text{HH}) = 1$ and $Q_2(\text{AD}) = 1$, respectively.

Definition 14. Let $\Delta \in \mathcal{C}_n$ and let $j \in \Delta$. The return map $\tilde{\varphi}_{[\Delta]} : S_n \rightarrow S_n$ is such that

$$\begin{aligned} \tilde{\varphi}_{[\Delta]}(x_0 x_1 \cdots x_j \cdots x_{n-1}, i_0 i_1 \cdots i_j \cdots i_{n-1}) \\ = (x_0 x_1 \cdots x_{j-1} x_{n-1}^* x_j \cdots x_{n-2}, \\ i_0 i_1 \cdots i_{j-1} i_{n-1} i_j \cdots i_{n-2}). \end{aligned} \quad (25)$$

That is, $\tilde{\varphi}_{[\Delta]}$ first inserts x_{n-1} and i_{n-1} to the position, where x_j and i_j are located, followed by changing x_{n-1} to the new value x_{n-1}^* . If $j, j' \in \Delta$ and $j < j'$ then $\tilde{\varphi}_{[\Delta]}$ is first applied to x_j and i_j .

Trivially, for any $n \geq 1$, $p \in S_n$, and $\Delta \in \mathcal{C}_n$,

$$\begin{aligned} \tilde{\varphi}_{[\Delta]}(\varphi_{[\Delta]}(p)) &= \varphi_{[\Delta]}(\tilde{\varphi}_{[\Delta]}(p)) \\ &= p. \end{aligned} \quad (26)$$

Theorem 15. Let $p \in S_n$. Define the image set and preimage set of p over $\mathcal{C}_n$ by

$$\begin{aligned} \mathcal{F}_p &= \bigcup_{\Delta \in \mathcal{C}_n} \{p' \in S_n : \varphi_{[\Delta]}(p) = p'\}, \\ \tilde{\mathcal{F}}_p &= \bigcup_{\Delta \in \mathcal{C}_n} \{\tilde{p} \in S_n : \varphi_{[\Delta]}(\tilde{p}) = p\}, \end{aligned} \quad (27)$$

respectively. For each $\Delta \in \mathcal{C}_n$,

- (a) there exists a unique $p' \in S_n$ such that $\varphi_{[\Delta]}(p) = p'$. Moreover, $|\mathcal{F}_p| = 2^n$;
- (b) there exists a unique $\tilde{p} \in S_n$ such that $\varphi_{[\Delta]}(\tilde{p}) = p$. Moreover, $|\tilde{\mathcal{F}}_p| = 2^n$.

Proof. (a) Let $x = x_0 x_1 \cdots x_{n-1}$ be the fitness state of $p \in S_n$. The image of x under both maps ϕ_Δ and ϕ_j is unique so p' is

unique. Note that $|\mathcal{C}_n| = 2^n$ and the set of images of x under $\varphi_{[\Delta]}$, $\Delta \in \mathcal{C}_n$, is the set of distinct binary numbers of length n , which can be associated with 2^n distinct states in S_n . This is a bijection between elements of $\mathcal{C}_n$ and the binary numbers of length n , so $|\mathcal{F}_p| = 2^n$.

(b) For $\Delta \in \mathcal{C}_n$, choose $\tilde{p} = \tilde{\varphi}_{[\Delta]}(p)$, where $\tilde{\varphi}_{[\Delta]}$ is the return map. The proofs of the uniqueness of $\tilde{p}$ and the cardinality of $\mathcal{F}_p$ follow a similar argument as in (a). $\square$

Theorem 16. Let $n \geq 2$ and $p \in S_n$. If $\Delta = \{n-m, n-1\}$, then $\varphi_{[\Delta]}^m(p) = p$ for $m = 2, 3, \dots, n$.

Proof. Fix n and m , where $2 \leq m \leq n$. Let $a = x_0 x_1 \dots x_{(n-m-1)}$ and $b = i_0 i_1 \dots i_{(n-m-1)}$ be sequences of length $(n-m)$ so that

$$p = (ax_{(n-m)}x_{(n-m+1)} \dots x_{(n-2)}x_{(n-1)}, \quad (28)$$

$$bi_{(n-m)}i_{(n-m+1)} \dots i_{(n-2)}i_{(n-1)}).$$

Let $\Delta = \{n-m, n-1\}$. Observe that, for any integer $\ell > 0$, a and b are fixed under $\varphi_{[\Delta]}^\ell$. Now

$$\begin{aligned} \varphi_{[\Delta]}(p) &= (ax_{(n-m+1)} \dots x_{(n-2)}x_{(n-1)}^*x_{(n-m)}^*, \\ &\quad bi_{(n-m+1)} \dots i_{(n-2)}i_{(n-1)}i_{(n-m)}) \\ \varphi_{[\Delta]}^2(p) &= (ax_{(n-m+2)} \dots x_{(n-2)}x_{(n-1)}^*x_{(n-m)} \\ &\quad \times x_{(n-m+1)}^*, bi_{(n-m+2)} \dots i_{(n-2)}i_{(n-1)} \\ &\quad \times i_{(n-m)}i_{(n-m+1)}) \\ &\vdots \\ \varphi_{[\Delta]}^\ell(p) &= (ax_{(n-m+\ell)} \dots x_{(n-2)}x_{(n-1)}^*x_{(n-m)} \dots x_{(n-m+\ell-2)} \\ &\quad \times x_{(n-m+\ell-1)}^*, bi_{(n-m+\ell)} \dots i_{(n-2)}i_{(n-1)} \\ &\quad \times i_{(n-m)} \dots i_{(n-m+\ell-2)}i_{(n-m+\ell-1)}) \end{aligned} \quad (29)$$

For $\ell = m-1$, $\varphi_{[\Delta]}^{m-1}(p) = (ax_{(n-1)}^*x_{n-m} \dots x_{n-2}^*, bi_{(n-1)}i_{n-m} \dots i_{n-2})$. Hence, $\varphi_{[\Delta]}(\varphi_{[\Delta]}^{m-1}(p)) = \varphi_{[\Delta]}^m(p) = p$. $\square$

Corollary 17. For any $n \geq 1$, $p \in S_n$, and $\Delta \in \{\{n-1\}, \{n-2, n-1\}, I_n\}$, the transition from p to $p' = \varphi_{[\Delta]}(p)$ is reversible.

The next theorem states some nonallowable transitions in S_n .

Theorem 18. Let

$$y = i_0 i_1 \dots i_{n-2} i_{n-1},$$

$$\bar{y} = i_{n-1} \bar{i}_1 \bar{i}_2 \dots \bar{i}_{n-2} \bar{i}_{n-1}, \quad \bar{i}_j \neq i_j \quad \text{for } j \in I_n \setminus \{0\}, \quad (30)$$

$$y' = i'_0 i'_1 \dots i'_{n-2} i'_{n-1}, \quad i'_j \neq i_j \quad \text{for } j \in I_n \setminus \{n-1\}.$$

Define

$$S_y = \{p \in S_n : y \text{ is the significance state of } p\},$$

$$S_{\bar{y}} = \{p \in S_n : \bar{y} \text{ is the significance state of } p\}, \quad (31)$$

$$S_{y'} = \{p \in S_n : y' \text{ is the significance state of } p\}.$$

- (a) Let $n \geq 2$. The only transitions in S_y , aside from self-transitions, are horizontal transitions. Let $n \geq 3$.
- (b) There is no transition from S_y to $S_{\bar{y}}$.
- (c) There is no transition between S_y and $S_{y'}$.

Proof. (a) It is easy to show that any transition between distinct points with the same significance state is under $\Delta = \{n-1\}$. In particular, horizontal transitions are between pairs $p = (x_p, y)$ and $\tilde{p} = (x_{\tilde{p}}, y)$, where $x_p = x_0 x_1 \dots x_{n-2} x_{n-1}$ and $x_{\tilde{p}} = x_0 x_1 \dots x_{n-2} x_{n-1}^*$.

(b) Let $p \in S_y$. The sets $\Delta = I_n$ and $\Delta = \{0, 1, \dots, n-2\}$ are the only elements of $\mathcal{C}_n$ such that i_{n-1} is moved to the leftmost position of the question order of $\varphi_{[\Delta]}(p)$, in which case question order must be $i_{n-1} i_{n-2} \dots i_0$, different from $\bar{y}$.

(c) As in (b), $\Delta = \emptyset$ and $\Delta = \{n-1\}$ are the only elements of $\mathcal{C}_n$ such that i_{n-1} remains in its position under $\varphi_{[\Delta]}$. The rest of the proof follows a similar argument as in (b). $\square$

2.5. Local Dynamics in S_n . Consider the subset

$$\mathcal{L}_m(\subset \mathcal{S}_n) = \{(c_0 c_1 \dots c_{(n-m-1)} x_{(n-m)} \dots x_{(n-1)},$$

$$\ell_0 \ell_1 \dots \ell_{n-m-1} i_{n-m} \dots i_{n-1}) \quad (32)$$

$$: c_j, \ell_j \text{ constant}, x_j \in \{0, 1\}, i_j \in I_n\},$$

where $n-m$ answers are constant. Since $|\mathcal{L}_m| = m!2^m$ states and $|\mathcal{S}_n| = n!2^n$ states, then S_n is composed of $2^{n-m}n!/m!$ subsets of the form $\mathcal{L}_m$. For $n-1$ constant answers (i.e., $m=1$), denote by $\mathcal{L}_{n-1}^{(c, \ell)}$ the subset of S_n , where question ℓ has constant answer c . Then any $p \in \mathcal{L}_{n-1}^{(c, \ell)}$ is given by $p = (cx_1 x_2 \dots x_{n-1}, \ell i_1 i_2 \dots i_{n-1})$, $x_j \in \{0, 1\}$, $i_j \in I_n$. Define $\mathcal{L}_{n-1}^{(c)}$ by

$$\mathcal{L}_{n-1}^{(c)} = \bigcup_{\ell \in I_n} \mathcal{L}_{n-1}^{(c, \ell)}. \quad (33)$$

Using (33), we can express S_n as

$$\mathcal{S}_n = \mathcal{L}_{n-1}^{(0)} \cup \mathcal{L}_{n-1}^{(1)}, \quad n \geq 1$$

$$= \mathcal{L}_n. \quad (34)$$

3. General Equation of Motion of Dynamics in S_n

This section concerns properties of all possible paths in S_n .

Definition 21 (see [13]). Let X and Y be arbitrary sets. A multivalued map F from X to Y , denoted by $F : X \rightrightarrows Y$, is such that $F(x)$ is assigned a set $Y_x \subset Y$ for all $x \in X$.

Consider our state space S_n . Denote by $\varphi_{[\Delta]} : S_n \rightarrow S_n$ the map in (22), where Δ_t is constant. Define

$$G = \{\varphi_{[\Delta]} : \Delta \in \mathcal{C}_n\} \quad (38)$$

and the multivalued map $F : S_n \rightrightarrows S_n$ by

$$F(p) = \bigcup \{g(p) \mid g \in G\}. \quad (39)$$

$|F(p)| = 2^n$ so F is a 1 to 2^n map. Under F , the orbit $\{p_t^k\}_{t \in \mathbb{N}_0}$ is such that $p_{t+1}^k \in F(p_t^k)$. The multivalued map F can be interpreted as a digraph whose $N = n!2^n$ vertices are the points in S_n and edge $p \rightarrow p'$ if $p' \in F(p)$.

Equivalently, F can be defined as a square matrix of size N that encodes all possible paths in S_n .

Definition 22. Let $V = \{1, 2, \dots, N\}$ and let $i, j \in V$. For $p, p' \in S_n$, let $\psi(p) = i$ and $\psi(p') = j$. Let F be the multivalued map in (39). The *theoretical transition matrix* for F under n questions is denoted by $T_n^{(th)} = (T_{n,ij}^{(th)})$, where

$$T_{n,ij}^{(th)} = \begin{cases} 1 & \text{if } p' \in F(p) \\ 0 & \text{otherwise.} \end{cases} \quad (40)$$

$T_{n,ij}^{(th)}$ captures all theoretically admissible transitions between states in S_n . Each entry $T_{n,ij}^{(th)} = 1$ indicates that there is a transition from state i to j in one step. Hence, $T_{n,ij}^{(th)}$ gives all possible paths in S_n that any observation can take.

Example 23. From Figure 4, we have $T_n^{(th)}$ for $n = 1$ and $n = 2$, as given below

$$T_1^{(th)} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad T_2^{(th)} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} \end{bmatrix}. \quad (41)$$

Because the digraph in S_1 is the same as the digraphs in $\mathcal{L}_1^{(c_j, \ell_j)} \subset S_2$ (Example 19), the matrix encoding transitions in S_1 is the same as the matrix encoding transitions in $\mathcal{L}_1^{(c_j, \ell_j)}$. The 2×2 submatrices of entries all equal to 1 (in bold numbers) in $T_2^{(th)}$ correspond to transitions in $\mathcal{L}_1^{(c_j, \ell_j)}$. The 2×2 zero submatrices in $T_2^{(th)}$ (in bold) denote the nonallowable transitions between states with the same significance (Theorem 18(a)).

3.1. Properties of $T_n^{(th)}$. We give some elementary results for $T_n^{(th)}$.

Theorem 24. (a) The digraph $\mathcal{G}_{T_n^{(th)}}$ is 2^n -regular; that is, the in- and out-degree of all its vertices is 2^n .

(b) The largest eigenvalue of $T_n^{(th)}$ is $\lambda_{\max}(T_n^{(th)}) = 2^n$, with associated eigenvector $v_{\lambda_{\max}} = \mathbf{1}$.

(c) Consider trace $\text{tr}(T_n^{(th)}) = N$.

(d) Consider $\det(T_n^{(th)}) = 0$.

Proof. (a) This is a consequence of Theorem 15.

(b) This is a consequence of (a) and the Perron-Frobenius theorem [14].

(c) Let $\psi(p) = i$. From the definition of self-transition, $\varphi_{[0]}(p) = p$ for any $p \in S_n$. Hence, $T_{n,ii}^{(th)} = 1$ for $i = 1, 2, \dots, N$.

(d) Let $p = (x, y)$ and $p' = (x', y)$ be points in S_n such that

$$x = x_0 x_1 \cdots x_{n-2} x_{n-1},$$

$$y = i_0 i_1 \cdots i_{n-2} i_{n-1},$$

$$x' = x_0 x_1 \cdots x_{n-2} x_{n-1}^*, \quad \text{where } x_{n-1}^* \text{ is given in (16).} \quad (42)$$

Claim. $\mathcal{F}_p = \mathcal{F}_{p'}$, where $\mathcal{F}_p$ and $\mathcal{F}_{p'}$ are given in (27).

The case where $n = 1$ is given in Example 23. We prove our claim for $n = 2$. By definition of self-transition, $\varphi_{[0]}(p) = p$ and $\varphi_{[0]}(p') = p'$. It is clear that $\varphi_{[\{1\}]}(p) = p'$ and $\varphi_{[\{1\}]}(p') = p$. Let

$$\begin{aligned} \varphi_{[\{0\}]}(p) &= p_1, \\ \varphi_{[\{0,1\}]}(p) &= p_2. \end{aligned} \quad (43)$$

Observe that

$$\begin{aligned} \varphi_{[\{0\}]}(p') &= \varphi_{[\{0\}]}(\varphi_{[\{1\}]}(p)) = \varphi_{[0,1]}(p) = p_2, \\ \varphi_{[\{0,1\}]}(p') &= \varphi_{[\{0,1\}]}(\varphi_{[\{1\}]}(p)) \\ &= \varphi_{[\{0\}]}(\varphi_{[\{1\}]}^2(p)) = \varphi_{[\{0\}]}(p) = p_1. \end{aligned} \quad (44)$$

Since p and p' have the same image for all $\Delta \in \mathcal{C}_2$, $\mathcal{F}_p = \mathcal{F}_{p'}$, as claimed.

Our claim implies that rows of $T_2^{(th)}$ associated with p and p' are the same. Because of these repeating rows, $\det(T_2^{(th)}) = 0$. The proof can be readily extended to the general case and is treated in exactly the same manner. $\square$

3.2. Subshift of Finite Type. Let $w_i, j \in V = \{1, 2, \dots, N\}$. If observation k is in state $s_j \in S_n$ after t time steps, we will write $p_t^k \in s_j$ to indicate that more than one observation may be in s_j . We associate with the orbit $\mathcal{O}_\varphi(k)$ a sequence of symbols

$w = (w_0 w_1 \cdots w_i \cdots)$, where $w_i = j$ if $p_i^k \in s_j$. Denote the corresponding *symbol space of one-sided sequences* of $T_n^{(\text{th})}$ by

$$\Sigma_{T_n^{(\text{th})}}^+ = \{w = (w_0 w_1 \cdots) \mid T_{n, w_i w_{i+1}}^{(\text{th})} = 1 \quad \forall i \in \mathbb{N}_0, w_i \in V\}. \quad (45)$$

The *equation of motion* in S_n is given by the shift map $\Sigma_{T_n^{(\text{th})}} : \Sigma_{T_n^{(\text{th})}}^+ \rightarrow \Sigma_{T_n^{(\text{th})}}^+$ defined by

$$\sigma_{T_n^{(\text{th})}}(w) = w' \quad \text{where } w'_i = w_{i+1} \quad (\forall i \geq 0). \quad (46)$$

Definition 25 (see [15]). The pair $(\sigma_{T_n^{(\text{th})}}, \Sigma_{T_n^{(\text{th})}}^+)$ is called the *subshift of finite type (SFT) determined by* $T_n^{(\text{th})}$.

The SFT determined by $T_n^{(\text{th})}$ captures the exact detail of all possible itineraries of observations in S_n . Many dynamical properties of an SFT depend on the structure of its associated transition matrix or digraph.

Definition 26 (see [16]). (i) A transition matrix T is *irreducible* if for each pair (i, j) , there exists $\ell > 0$ so that $T^\ell > 0$. Otherwise, T is *reducible*. (ii) The digraph $G_T = (V, E)$ associated with T is *strongly connected* if it is possible to get from any vertex to any other one by traversing a sequence of edges as directed by T . (iii) If there exists $\ell' > 0$ such that $T^\ell > 0$ for all $\ell > \ell'$, then T is *primitive*. (iv) Let X be any set. A continuous map $f : X \rightarrow X$ is (*topologically*) *transitive* if, for all nonempty open sets $U, U' \subset X$, there exists $\ell \geq 0$ such that $f^\ell(U) \cap U' \neq \emptyset$. (v) If there is $\ell' > 0$ such that $f^\ell(U) \cap U' \neq \emptyset$ for all $\ell \geq \ell'$, then f is (*topologically*) *mixing*.

Consider the SFT given by (σ_T, Σ_T^+) . The following results [17, 18] concern dynamical properties of a transitive SFT and algebraic properties of its transition matrix. (i) The shift map Σ_T is transitive (resp., mixing) if and only if T is irreducible (resp., primitive). (ii) T is irreducible if and only if the digraph $G_T = (V, E)$ is strongly connected. In that case we say that V is irreducible. (iii) A nonnegative, irreducible matrix with a positive element on the main diagonal is primitive.

Theorem 27. $T_n^{(\text{th})}$ is irreducible.

Proof. We prove by induction on n . From Example 23, $T_{1,ij}^{(\text{th})} = 1$ for all (i, j) so $T_n^{(\text{th})}$ is irreducible for $n = 1$. Assume that $T_n^{(\text{th})}$ is irreducible for $n = m$. Then the digraph $G_{T_m^{(\text{th})}}$ is strongly connected, and S_m is irreducible. We prove that $G_{T_{m+1}^{(\text{th})}}$ is strongly connected to $n = m + 1$.

Recall that transitions in S_{m+1} are under the set $\mathcal{E}_{m+1}$. From (37), this set can be expressed as $\mathcal{E}_{m+1} = \mathcal{D}_m^{(m+1)} \cup (\mathcal{E}_{m+1} \setminus \mathcal{D}_m^{(m+1)})$, where $\mathcal{D}_m^{(m+1)}$ is the set associated with transitions in the irreducible set $\mathcal{L}_m^{(c,i_j)} (= \mathcal{S}_m)$, $c \in \{0, 1\}$ and $i_j \in I_{m+1}$. With the additional transitions under $\mathcal{E}_{m+1} \setminus \mathcal{D}_m^{(m+1)}$, we show that

(i) the set $\mathcal{L}_m^{(c)} \subset \mathcal{S}_{m+1}$ (given in (33)) is irreducible;

(ii) there is a path to and from any pair p, p' , where $p \in \mathcal{L}_m^{(0)}$ and $p' \in \mathcal{L}_m^{(1)}$.

We prove (i) by showing that there is $p \in \mathcal{L}_m^{(c,i_j)}$ such that, for all $i'_j \neq i_j$, there is a transition from p to $p' \in \mathcal{L}_m^{(c,i'_j)}$. By Theorem 16, $\varphi_{[[0,m]]}^{m+1}(p) = p$ for any $p \in \mathcal{S}_{m+1}$. Let

$$\begin{aligned} p_0 &= (cc \cdots cc^*, i_0 i_1 \cdots i_{m-1} i_m) \\ p_1 &= \varphi_{[[0,m]]}(p_0) = (cc \cdots cc^*, i_1 i_2 \cdots i_m i_0) \\ p_2 &= \varphi_{[[0,m]]}^2(p_1) = (cc \cdots cc^*, i_2 i_3 \cdots i_0 i_1) \\ &\vdots \\ p_m &= \varphi_{[[0,m]]}^m(p_{m-1}) = (cc \cdots cc^*, i_m i_0 \cdots i_{m-2} i_{m-1}). \end{aligned} \quad (47)$$

Observe that all p_j 's have the same fitness states but distinct significance states. In particular, each p_j is contained in a distinct subset $\mathcal{L}_m^{(c,i_j)}$. This vertical transition allows all $p_j \in \mathcal{L}_m^{(c,i_j)}$ to visit a distinct $\mathcal{L}_m^{(c,i'_j)}$ in at most m steps. Hence, $\mathcal{L}_m^{(c)}$ is irreducible. To show (ii), take

$$\begin{aligned} p &= (00 \cdots 0, m(m-1) \cdots 10) \in \mathcal{L}_m^{(0)}, \\ p' &= (11 \cdots 1, 01 \cdots (m-1)m) \in \mathcal{L}_m^{(1)}. \end{aligned} \quad (48)$$

Under $\Delta = I_{m+1}$, there is a reversible transition $p \xleftrightarrow{[\Delta]} p'$.

From (i), (ii), and the irreducibility of $T_m^{(\text{th})}$, there is a path between any pair $p, p' \in \mathcal{S}_{m+1}$. Hence, $G_{T_{m+1}^{(\text{th})}}$ is strongly connected and $T_{m+1}^{(\text{th})}$ is irreducible, as desired. $\square$

Corollary 28. $T_n^{(\text{th})}$ is primitive.

Definition 29 (see [15]). The topological entropy of the shift map $\Sigma_T : \Sigma_T^+ \rightarrow \Sigma_T^+$ is defined by

$$h_{\text{top}}(\sigma_T) = \lim_{m \rightarrow \infty} \frac{\ln(|\mathcal{W}(m)|)}{m}, \quad (49)$$

where $\mathcal{W}(m)$ is the set of allowable sequences of length $m \geq 1$.

A continuous map f on a compact metric space X is chaotic if $h_{\text{top}}(f) > 0$ [17]. The topological entropy for SFTs is given by the following theorem [15].

Theorem 30. Let T be a transition matrix and let $\Sigma_T : \Sigma_T^+ \rightarrow \Sigma_T^+$ be the associated SFT. Then

$$h_{\text{top}}(\sigma_T) = \ln(\lambda_{\max}), \quad (50)$$

where $\lambda_{\max}$ is the maximum eigenvalue of T .

FIGURE 6: (a) Orbits of Agincourt households in S_3 , over time. (b) Number of orbit visits to each state in S_3 , where majority are in $\mathcal{L}$, having the favourable property of being headed by an adult (i.e., $Q_1 = 1$).

Since $(\sigma_{T_n^{(th)}}, \Sigma_{T_n^{(th)}}^+)$ is the associated subshift of the multi-valued map F in (39), one has

$$\begin{aligned} h_{\text{top}}(F) &= h_{\text{top}}(\sigma_{T_n^{(th)}}) \\ &= \ln(2^n) \\ &= n \ln 2 > 0. \end{aligned} \quad (51)$$

4. Dynamics from Data

Given multivariable longitudinal binary data of dimension n , every observed orbit is an orbit of an SFT determined by $T_n^{(th)}$. Dynamics of real-world longitudinal data however is not often defined by an SFT. Data usually selects certain paths given by $T_n^{(th)}$ and may sometimes stay in a particular subspace of S_n .

4.1. Nonstationary SFT. Let $p, p' \in S_n$, $\psi(p) = i$, and $\psi(p') = j$. For each $t \geq 0$, denote by $T_{n,t}^{(data)}$ the matrix that records observed transitions that occur in the longitudinal data from time t to $t + 1$, where

$$(T_{n,t}^{(data)})_{ij} = \begin{cases} 1 & \text{if } \exists k \in K, \text{ such that } p \xrightarrow{[\Delta_i^k]} p' \\ 0 & \text{otherwise.} \end{cases} \quad (52)$$

We note that $T_{n,t}^{(data)}$ is defined over an interval of time and that $T_{n,t}^{(data)}$ may vary with time. Some allowable transitions between states given by $T_n^{(th)}$ might not occur in the observed data so $T_{n,t}^{(data)} \neq T_n^{(th)}$. For the case where $T_{n,t}^{(data)}$ is constant we write $T_{n,t}^{(data)} \equiv T_n^{(data)}$ and we can define the SFT given by the pair $(\sigma_{T_n^{(data)}}, \Sigma_{T_n^{(data)}}^+)$.

If $T_{n,t}^{(data)}$ is not constant, then we have a sequence of matrices

$$A = \{T_{n,t}^{(data)}\}_{t \geq 0}. \quad (53)$$

Given A , define

$$\begin{aligned} \Sigma_A^{(data)} &= \{w = (w_0 w_1 \cdots w_{\mathcal{T}} \cdots) \in \Sigma_{T_n^{(th)}}^+ \\ &\quad : (T_{n,t}^{(data)})_{w_t w_{t+1}} = 1\}. \end{aligned} \quad (54)$$

We call $\Sigma_A^{(data)}$ the *nonstationary symbolic space restricted by the sequence of matrices A* . The shift takes place as usual in $\Sigma_A^{(data)}$ and is denoted by $\sigma_A : \Sigma_A^{(data)} \rightarrow \Sigma_A^{(data)}$. We refer the reader to [19–21] for a discussion on nonstationary SFT.

Definition 31. The pair $(\sigma_A, \Sigma_A^{(data)})$ is the nonstationary SFT (NSFT) determined by the sequence of matrices A .

Visualization of dynamics of longitudinal data defined by an NSFT is illustrated by a sequence of directed graphs called a Bratteli diagram [9, 10]. Equivalently we may plot orbits of observations in S_n over time. Figure 6(a) illustrates the orbits of a population in a subset of S_3 over time.

4.2. Population Dynamics. We discuss the longitudinal data of a population of $K > 1$ observations. Aside from all possible paths that an observation can take in S_n , we are also interested in the number of observations on paths. Because it is possible for more than one observation to occupy a state in S_n at a given time, we can consider the number of observations that follow the same transition in S_n .

In general, given a transition matrix T (e.g., those in (40) or (52)), standard analysis is to accumulate number of transitions between states, and from this construct the associated stochastic transition matrix. Construction of associated

transition and stochastic matrices on states of S_n can then follow as usual. In what follows, let V denote a finite index set with $|V| \geq 2$ and $i, j \in V$.

Definition 32. Let $H_t(i)$ be the number of observations in state i at time t . The density matrix at time t is defined by

$$D_t = (d_{t,ij}), \quad (55)$$

where $d_{t,ij}$ is the number of observations in state i at time t that go to state j at time $t + 1$. The (net) flux of state i at time t is defined by

$$F_t(i) = H_{t+1}(i) - H_t(i). \quad (56)$$

Let $\ell_{t,ij} = d_{t,ji} - d_{t,ij}$, $i, j \in V$. The flux in (56) can also be expressed as

$$F_t(i) = \sum_j \ell_{t,ij}. \quad (57)$$

From (56) and (57), we have

$$H_t(i) = \sum_j d_{t,ij}, \quad H_{t+1}(i) = \sum_j d_{t,ji}. \quad (58)$$

Let H_t be a row vector whose i th entry is $H_t(i)$. We will refer to H_t as the *observed capacity vector* at time t . Given the initial observed capacity vector H_0 , there are two methods that we can use to determine H_{t+1} .

(i) *Nonhomogenous Case.* From data, we are encouraged to construct a probability matrix based directly on the density matrices D_t . For each D_t , we construct a time-dependent probability matrix $P_t = (P_{t,ij})$, where

$$P_{t,ij} = \begin{cases} \frac{d_{t,ij}}{\sum_j d_{t,ij}}, & \text{if } \sum_j d_{t,ij} \neq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (59)$$

The capacity vector H_{t+1} is given by the product

$$\begin{aligned} H_{t+1} &= H_t \cdot P_t \\ &= H_0 \cdot \prod_{m=0}^t P_m. \end{aligned} \quad (60)$$

We show that the capacity $H_{t+1}(i)$ in (56) agrees with the i th entry of (60). It is trivial if $\sum_j d_{t,ij} = 0$ for all t . Otherwise we have

$$\begin{aligned} H_{t+1}(i) &= H_t(i) + F_t(i) \\ &= H_t(i) + \sum_j (d_{t,ji} - d_{t,ij}) \\ &= H_t(i) + \sum_j \left(P_{t,ji} \sum_i d_{t,ji} - P_{t,ij} \sum_j d_{t,ij} \right) \end{aligned}$$

$$\begin{aligned} &= H_t(i) + \sum_j P_{t,ji} H_t(j) - \sum_j P_{t,ij} H_t(i) \\ &= H_t(i) \left(1 - \sum_j P_{t,ij} \right) + \sum_j P_{t,ji} H_t(j) \\ &= \sum_j P_{t,ji} H_t(j). \end{aligned} \quad (61)$$

(ii) *Homogenous Case.* Let D_t be the density matrix at time t and let

$$D^{(\text{data})} = \sum_{t=0}^{T-1} D_t \quad (62)$$

be the accumulated density matrix of the data over the observation period. Define the mean density matrix by

$$D = \frac{1}{T} D^{(\text{data})}. \quad (63)$$

Suppose D is irreducible. Define the mean probability matrix from D by $P_D = (P_{D,ij})$, where

$$P_{D,ij} = \frac{D_{ij}}{\sum_j D_{ij}}. \quad (64)$$

The capacity vector at $t + 1$ is given by the product

$$\begin{aligned} H_{t+1} &= H_t \cdot (P_D) \\ &= H_0 \cdot (P_D)^t. \end{aligned} \quad (65)$$

The probabilities in (64) are based on observed frequencies at which one state follows another. The transition probabilities of a Markov model using states in S_n are then estimated. Markov modeling on the bigger space S_n may give better results in generating data as S_n is a more detailed space than a space where order of variable is fixed (e.g., M_n).

Definition 33. A set of observations K' cluster in S_n if for all $k \in K'$, there exists $R \subset S_n$ and time interval I such that the orbit $O(k) \subset R$ for all $t \in I$.

An advantage in using the bigger space S_n to analyse multivariable longitudinal data is that all possible clusterings in stable variables are revealed. We can also visualize in S_n the shift in clusters from one region R to another region R' . There is much interesting information that we can gather from a cluster in a subset R of S_n . Clustering in $\mathcal{L}_m$ ($m \neq n$) defined in (32) involves observations with the same stable $n - m$ variables. Subsets $\mathcal{L}_m$ are important as they automatically determine which variables are trivial in the analysis. We either reduce analysis from S_n to $\mathcal{L}_m$ or replace the $n - m$ constant variables with other nonconstant variables that are of interest to purpose. In addition, orbits that spend most (if not all) of the time in L_m have strong correlation among the $n - m$ variables.

TABLE 5: Agincourt questionnaire with corresponding coded answer and frequency of answer change in the population.

Question	Answer	Code	f_i^P
Q_0 : Is the biological mother present (>6 months)? (BM)	Yes	1	760
Q_1 : Is the household headed by a minor? (HH)	Yes	0	4
Q_2 : Is there an adult death in the household? (AD)	Yes	0	119

TABLE 6: Total transitions from state $i \rightarrow j$ ($i, j = 23, 24$) in defaulting and nondefaulting household, together with the odds ratios (95% confidence interval and P value).

State $i \rightarrow j$	Defaulting	Nondefaulting	Odds ratio (95% CI; P value)
23 $\rightarrow$ 23	256	63	1.59 (1.17 to 2.17; $P = 0.0031$)
23 $\rightarrow$ 24	238	84	1.00 (0.75 to 1.33; $P = 0.9925$)
24 $\rightarrow$ 23	260	96	0.94 (0.72 to 1.24; $P = 0.6653$)
24 $\rightarrow$ 24	214	99	0.70 (0.53 to 0.92; $P = 0.0110$)

5. Application

We illustrate our method using the $n = 3$ demographic binary questions taken from Agincourt [22]. With purpose

$\mathcal{P}$: to investigate the effect of household change on a child's educational progress, the three questions in (4) are associated with household change and are all of interest to $\mathcal{P}$. Table 5 gives the question (indexed by 0, 1, and 2, resp.) with associated variables (BM, HH, and AD), favourable/unfavourable coding, and frequency of change.

The longitudinal data consists of $K = 188$ observations with no missing data. Each household is surveyed each year, for nine years [22], and the population is constant throughout the observation period; that is, no households enter nor leave the population. Using (20), the frequency relation in the population P is $f_1^P \ll f_2^P \ll f_0^P$. The number of unfavourable = 0 answers (over 1692 answers) is Q_0 : 844(49.9%), Q_1 : 2(0.1%), and Q_2 : 71(4.2%). We expect that this, together with the very small value of f_1^P , will give clustering in the question order 1^{**} , where $Q_1 = 1$.

Figure 6(a) illustrates orbits of households in S_3 over time, where a red edge denotes a transition to the left, a green edge is a transition to the right, and a blue edge is a self-transition. Most of the orbits are in

$$\begin{aligned} \tilde{\mathcal{L}} &= \{21, 22, 23, 24, 29, 30, 31, 32\} \\ &= \{(x, y) : x = 1^{**}, y = 1^{**}\} \end{aligned} \quad (66)$$

and are characterised by the favourable property of being headed by an adult (i.e., $Q_1 = 1$). Of the 179 households that stay in $\tilde{\mathcal{L}}$, 159 distinct orbits are found. However, we see from Figure 6(b) that the number of state visits in S_3 is dominant in states 23 = (110, 120) and 24 = (111, 120) so dynamics principally takes place in these two states.

State 23 is characterised by the most stable variable HH = 1, followed by AD = 1 and then by BM = 0. Similarly, state 24 is characterised by the most stable variable HH = 1, followed by AD = 1 and then by BM = 1. Orbits that stay strictly in states 23 and 24 are associated with households headed by an adult and with no adult death during the survey period. Those idle in 23 (resp., 24) have mothers away (resp., present) during the survey period. Households that transition from 23 to 24 are characterized by mother out- to in-migration, while transitions from 24 to 23 denote in- to out-migration.

With regard to purpose, we define educational default by the number of years that a child has failed in his school life. A household with a child who has failed 4 years or more in his school life is classified as defaulting. Orbits in Figure 6(a) are then split into defaulting and nondefaulting populations, as shown in Figures 7(a) and 7(b), respectively.

Note that if we include a fourth question

$$Q_3: \text{is the household defaulting (=0) or nondefaulting (=1)?} \quad (67)$$

then orbits will be plotted in S_4 , as seen in Figure 5. Orbits that cluster on the left are the defaulting households ($Q_4 = 0$), while those on the right are nondefaulting ($Q_4 = 1$).

Figure 8 gives $H_t(i)$, the number of observed defaulting and nondefaulting households at each state $i = 23, 24$ from $t = 1999$ to $t = 2007$. It is clear that, for both subpopulations, there is an exchange of numbers between states 23 (BM-out) and 24 (BM-in). Figures 9(a) and 9(b) illustrate the orbits of defaulting and nondefaulting households over time, and the accumulated number of visits in S_4 , respectively.

Table 6 gives the (accumulated) number of state transitions $i \rightarrow j$, $i, j \in \{23, 24\}$ for both subpopulations. Using odds ratio, we have the value 1.59 (>1) for transition 23 $\rightarrow$ 23, which means that once mothers from defaulting households out-migrate, they take long to return to their households. Out-migrating mothers are then more likely to be away from home in the defaulting households.

FIGURE 7: Orbits in S_3 of (a) defaulting and (b) nondefaulting households. FIGURE 8: Number of defaulting and nondefaulting households in states 23 and 24 for $t = 1999$ to $t = 2007$.

For $24 \rightarrow 24$ we have odds ratio of 0.70 (< 1) indicating that in-migrating mothers are more likely to stay at home in the nondefaulting households. Both odds ratios have significant P values and these support our hypothesis that mother out-migration places child educational progression at risk. We have no strong evidence that mother out-migration ($24 \rightarrow 23$) is more likely in the defaulting households nor that in-migration ($23 \rightarrow 24$) is more likely in nondefaulting households.

6. Concluding Remarks

By defining a specific swap operation on multivariable longitudinal data, we have imposed dynamics on data. No approximations are made and orbit of an observation in the 2-dimensional space S_n can be decoded back to

the observation's original data. By including the order of variables in our analysis, we have introduced new states, along with an objective criterion of relative fitness and identification of subshift. We have only considered binary-valued variables but the method may be generalized to m -ary valued multivariable longitudinal data. Although the method of orbits is nonstatistical, it can however be used to aid in statistical analysis.

The advantage in plotting orbits in S_n is that information of change can be extracted directly from the visualized orbits. In addition, orbits in S_n facilitate the identification of cause and effect, of preceding events that might uniquely and usefully associate with change. Frequency of change is a property taken directly from the data so that, under our reordering operation, any practitioner with the same favourable and unfavourable coding will find the same ordering of questions and arrive at the same analysis of longitudinal data.

FIGURE 9: (a) Orbits of defaulting (left cluster) and nondefaulting (right cluster) households in S_4 , over time. (b) Accumulated number of visits to each state of S_4 .

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

Maria Vivien Visaya is supported by a Claude Leon Fellowship.

References

- [1] E. S. Fung, W. K. Ching, S. Chu, M. K. Ng, and W. Zang, "Multivariate Markov chain models," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, vol. 3, pp. 298–302, Hammamet, Tunisia, October 2002.
- [2] C. Bijleveld, L. J. T. van der Kamp, A. Mooijjaart, W. van der Kloot, R. van der Leeden, and E. van der Burg, *Longitudinal Data Analysis: Designs, Models and Methods*, Sage, Thousand Oaks, Calif, USA, 1998.

- [3] H. Verbeke, S. Fiews, G. Molenberghs, and M. Davidian, "The analysis of multivariate longitudinal data: a review," *Statistical Methods in Medical Research*, 2012.
- [4] I. Vlachos and D. Kugiumtzis, "State space reconstruction for multivariate time series prediction," *Nonlinear Phenomena in Complex Systems*, vol. 11, no. 2, pp. 241–249, 2008.
- [5] J. Al-Aziz, N. Christou, and I. D. Dinov, "SOCR motion charts: an efficient, open-source, interactive and dynamic applet for visualizing longitudinal multivariate data," *Journal of Statistics Education*, vol. 18, no. 3, pp. 1–29, 2010.
- [6] A. Grossenbacher, "The globalisation of statistical content," *Statistical Journal of the IAOS*, vol. 25, no. 3-4, pp. 133–144, 2008.
- [7] A. Inselberg, *Parallel Coordinates: Visual Multidimensional Geometry and Its Applications*, Springer, New York, NY, USA, 2009.
- [8] H. Zhou, X. Yuan, H. Qu, W. Cui, and B. Chen, "Visual clustering in parallel coordinates," in *Proceedings of the Eurographics/IEEE-VGTC Symposium on Visualization*, A. Vilanova, A. Telea, G. Scheuermann, and T. Möller, Eds., vol. 27, 2008.
- [9] V. Basios, G. Forti, and G. Nicolis, "Symbolic dynamics generated by a combination of graphs," *International Journal of Bifurcation and Chaos*, vol. 18, no. 8, pp. 2265–2274, 2008.
- [10] F. Durand, "Combinatorics on Bratteli diagrams and dynamical systems," in *Combinatorics, Automata and Number Theory*, V. Berth and M. Rigo, Eds., pp. 324–372, 2010.
- [11] L. Beckett and P. Diaconis, "Spectral analysis for discrete longitudinal data," *Advances in Mathematics*, vol. 103, no. 1, pp. 107–128, 1994.
- [12] A. Gottschau, "Markov chain models for multivariate binary panel data," *Scandinavian Journal of Statistics*, vol. 21, no. 1, pp. 57–71, 1994.
- [13] M. Mrozek, "Inheritable properties and computer assisted proofs in dynamics," in *Scientific Computing and Validated Numerics*, G. Alefeld, A. Frommer, and B. Lang, Eds., vol. 90, pp. 245–257, Akademie, Berlin, Germany, 1996.
- [14] H. Minc, *Nonnegative Matrices*, John Wiley & Sons, New York, NY, USA, 1998.
- [15] C. Robinson, *Dynamical Systems: Stability, Symbolic Dynamics and Chaos*, CRC Press, New York, NY, USA, 2nd edition, 1999.
- [16] B. Kitchens, *Symbolic Dynamics: One-Sided, Two-Sided and Countable State Markov Chains*, Springer, New York, NY, USA, 1998.
- [17] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, New York, NY, USA, 1995.
- [18] R. Varga, *Matrix Iterative Analysis*, Prentice Hall, Englewood Cliffs, NJ, USA, 1962.
- [19] P. Arnoux and A. M. Fisher, "Anosov families, renormalization and non-stationary subshifts," *Ergodic Theory and Dynamical Systems*, vol. 25, no. 3, pp. 661–709, 2005.
- [20] A. H. Fan and M. Pollicott, "Non-homogeneous equilibrium states and convergence speeds of averaging operators," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 129, no. 1, pp. 99–115, 2000.
- [21] A. M. Fisher, "Nonstationary mixing and the unique ergodicity of adic transformations," *Stochastics and Dynamics*, vol. 9, no. 3, pp. 335–391, 2009.
- [22] K. Kahn, S. M. Tollman, M. A. Collinson et al., "Research into health, population and social transitions in rural South Africa: data and methods of the Agincourt health and demographic surveillance system," *Scandinavian Journal of Public Health*, vol. 35, supplement 69, pp. 8–20, 2007.