A Comparison of Alternative Methods for Describing Life Course Trajectories

Alberto Palloni
James M. Raymo
Department of Sociology
University of Wisconsin-Madison

John R. Warren
Andrew Halpern-Manners
Liying Luo
Minnesota Population Center
Department of Sociology
University of Minnesota

Version: September 2011

Paper submitted for prepared for presentation at the annual meetings of the Population Association of America, May 2012, San Francisco. This research uses data from the Wisconsin Longitudinal Study (WLS) of the University of Wisconsin-Madison. Since 1991, the WLS has been supported principally by the National Institute on Aging (AG-9775 and AG-21079), with additional support from the Vilas Estate Trust, the National Science Foundation, the Spencer Foundation, and the Graduate School of the University of Wisconsin-Madison. A public use file of data from the Wisconsin Longitudinal Study is available from the Data and Program Library Service, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, Wisconsin 53706 and at http://dpls.dacc.wisc.edu/WLS/wlsarch.htm. All errors omissions, however, are the responsibility of the authors. Please direct correspondence to John Robert Warren, 909 Social Sciences, 267 19th Ave S, Minneapolis, MN 55455 or email ware046@umn.edu.

A Comparison of Alternative Methods for Describing Life Course Trajectories

ABSTRACT

Social science researchers have increasing access to rich data that include repeated measures of individuals' attributes at narrow age intervals over long periods of the life course. Data on age-graded "trajectories" of attributes are being used to test theories in population studies, sociology, criminology, public health, and beyond. Despite the richness of modern "trajectory" data, there is no consensus about which technique to use to (1) identify the number of trajectories in a population; (2) describe the attributes of those trajectories; or (3) decide which trajectories best describe each individual's biography. We will model simulated trajectory data on women's employment statuses across their careers. Do commonly used methods (i.e., group-based trajectory modeling, optimal matching analysis, grade of membership analysis, growth mixture modeling, and naïve classification) yield the same results? If not, under what circumstances do particular methods produce valid results?

A Comparison of Alternative Methods for Describing Life Course Trajectories

Social scientists studying any number of substantive issues have developed rich theories involving concepts that vary systematically with age. These theories explain variation with age in boys' delinquent behaviors, women's rates of paid employment, research professors' levels of productivity, and elementary school students' reading ability, to name just a few. In many of these areas of inquiries, scholars have utilized ideas about *trajectories* of individual-level attributes, where a trajectory is an age-sequenced series of measurements that reflects career or life course patterns of stability and change in the focal attribute.

Until recently, the richness of theories involving trajectories has outpaced the quality of the data available to test those theories. Whereas theories about boys' delinquency, women's paid employment, students' reading abilities, and others imply the need for detailed data on these characteristics collected at short age intervals over long spans of people's lives, researchers have usually been left testing these theories using data collected at wide age intervals (typically one or more years) over decidedly truncated portions of the life course. For example, theories involving trajectories of elementary school students' reading ability suggest the need for data collected at multiple points each academic year (and over the summer) and from early adolescence through the teenage years. In fact, researchers looking to test these theories have typically had to rely on data collected annually (or less frequently) and for just a few years.

In recent years, however, computer and data storage technologies, new methods for collecting data, and other developments have combined to produce much richer age-graded trajectory data. Across numerous substantive domains, it is increasingly possible to utilize data collected at frequent intervals over longer spans of the life course or career. For example, various data sets now include annual (or more frequent) information about people's labor force status, earnings, and job conditions over the bulk of the working career, and other data sets include annual (or more frequent) measures of delinquent behaviors across the full span of adolescence. The availability of these data is making possible more complete and more useful tests a wide range of theories that involve age-graded trajectories of attributes or conditions.

However, the relatively recent availability of "trajectory data" — data on some attribute (or set of attributes) collected at short age intervals over long spans of the life course or career — has raised new methodological challenges. Across substantive domains, social scientists modeling trajectory data frequently find themselves asking similar questions: Is there a finite number of trajectories in a population, or are individuals' biographies all essentially unique? If there is a finite number of trajectories, how should researchers determine what that number is? Once researchers have determined how many trajectories there are in a population, how should they determine which trajectory best describes particular individuals' biographies? (Or, put another way, how should they determine which individuals have experienced categorically similar trajectories of outcomes or attributes?) How can researchers quantify their uncertainty with respect to how many trajectories there are in a population, or which trajectories best describe particular individuals' biographies?

There are now several techniques available for modeling trajectory data. One popular option is to dictate, *a priori*, that there is some fixed number of trajectories and to use some *ad hoc* algorithm to declare that particular individuals' biographies are described by particular trajectories; although widely used, such "naïve" methods do not provide falsifiable results, and cannot quantify uncertainty. On the other hand, group-based trajectory modeling (e.g., Nagin 2005), optimal matching analysis (e.g., Abbott and Tsay 2000), grade of membership analysis (e.g., Manton 1991), and growth mixture modeling (e.g., Muthén 2004) are all more rigorous methods for defining trajectories, for declaring which trajectory best describes each individual's biography, and for quantifying uncertainty about these things. All of these methods are now routinely used in the social sciences (and beyond) to model trajectory data. However, researchers investigating specific empirical questions typically employ just *one* of these methods, with relatively little attention to whether the choice of method matters for substantive research findings.

Research Questions

The goal of our paper is to assess the relative utility of various methods for identifying and characterizing age-graded trajectories. When analyzing the same data, do the five methods listed above—group-based trajectory modeling, optimal matching analysis, grade of membership analysis, growth mixture modeling, and naïve classification—yield the same conclusions about (1) the finite number of age-graded trajectories that exist in the data; (2) the characteristics or qualities of those various trajectories; and (3) decisions about which particular trajectory best describes each individual's biography? If the five methods listed above do not yield the same information about these three features of trajectories, then under what circumstances do these methods produce different results? Perhaps most importantly, under what circumstances does each method produce valid results?

If a researcher's choice of method for modeling trajectory data makes little difference for their conclusions about the number or qualities of trajectories or about which trajectory best describes each individual's biography, then this will be reassuring: We will have greater confidence in the validity of any number of substantive scientific projects that have (more or less uncritically) elected to use just one from among these various methods. On the other hand, if the choice of method does matter for these conclusions, then it will be imperative to understand (1) the situations in which the choice of methods is most likely to matter and (2) which methods produce the most accurate results under particular conditions.

Research Plan

We will begin by generating a large number of simulated data sets that characterize women's employment statuses between ages 18 and 64. That is, for each (simulated) woman we will "observe" her employment status at frequent age intervals and across the bulk of her working career. Each woman's trajectory of employment statuses will be simulated using a random number generator; at each time point, her probability of employment will be modeled as a function of her age, education, marital status, and number of children at that point in time. Parameters used to inform this simulation and its data-generating equations will be based on inferences drawn from data that include life-course long measurements of women's employment status (such as the 1975 through 2010 waves of the Wisconsin Longitudinal Study). Once we have the simulated data set, we will use the five methods listed above (1) to estimate the number of trajectories of women's employment status in the population, (2) to describe or characterize the attributes of those trajectories, and (3) to determine which trajectory best describes each woman's own individual biography of employment statuses.

Although modeling one simulated data set will allow us to understand whether the five methods listed above yield the same results, it will not provide information about the conditions under which the methods produce different results or about the conditions under which particular methods produce valid or invalid results. Consequently, we will generate a series of simulated data sets that differ with respect to (1) how many trajectories actually characterize women's employment statuses and (2) the level of certainty with which women can be said to "belong" to particular trajectory groups. That is, the simulated data sets will vary in one dimension with respect to whether there are really k=1, 2, … *k* underlying trajectories that characterize women's experiences with paid employment. At the same time, the simulated data sets will vary along another dimension with respect to how rigidly women follow particular trajectories in their own lives. This strategy will allow us to assess the circumstances under which different methods produce different results. What is more, because the simulated data are all generated using known parameters, we can assess the conditions under which particular methods produce accurate and reliable results. Results of these analyses will provide an empirical basis for researchers to make informed decisions about the measurement of life trajectories using the rich longitudinal data to which we now have access.