Measuring Resemblance in Sequence Data: An Optimal Matching Analysis of Musicians' Careers[1]

Andrew Abbott
Rutgers University

Alexandra Hrycak
University of Chicago

This article introduces a method that measures resemblance between sequences using a simple metric based on the insertions, deletions, and substitutions required to transform one sequence into another. The method, called optimal matching, is widely used in natural science. The article reviews the literature on sequence analysis, then discusses the optimal matching algorithm in some detail. Applying this technique to a data set detailing careers of musicians active in Germany in the 18th century demonstrates the practical steps involved in the application of the technique and develops a set of typical careers that successfully categorize most of the actual careers studied by the authors.

Although sociologists have long studied careers, assessing resemblance among career patterns remains difficult. Some analysts have resorted to enumeration; others have turned to stochastic models. The central problem is a simple one; with even a few events measured at regular intervals, the number of possible careers rapidly surpasses calculability.

In this article we present a technique for the direct measurement of career resemblance. Known in many disciplines, this so-called optimal matching method applies to data consisting of ordered sequences of potentially repeating events. Although we illustrate the method using data on careers, it could be used to analyze the track of occupational profes-

[1] The original work for this article was done while the junior author was an undergraduate writing an honors thesis (Hrycak 1988) under the senior author's direction. Partial support for the work was provided by the Rutgers Research Council. We would like to thank several anonymous referees for demanding clarifications that have improved the article considerably. We would also like to thank Ron Angel for his commentary.

0002-9602/91/9601-0005$01.50

sionalization, the development of organizations, the life cycles of individuals or families, the unfolding of revolutions, or any other "history" that can be expressed as a sequential list of events. Since it makes no stochastic assumptions, it is also appropriate for sequences not generated step by step, such as sequences of steps in the labor process or of stages in a ritual.

The first half of the article—Sections I, II, and III—places the technique in its context. We first consider the substantive literature on careers, then turn to an overview of current sequence analysis methods. In Section III, we present optimal matching techniques themselves. The second part of the article—Sections IV, V, and VI—illustrates these techniques with data on careers of musicians active in Germany during the period 1660–1800. We first discuss the structure and organization of the labor market. We then analyze the careers of about 300 musicians, using optimal matching to seek typical career patterns. The article closes with a discussion of the strengths and weaknesses of the measures. The Appendix contains a detailed exposition of optimal matching and a discussion of computation and software issues.

I. THE CAREERS LITERATURE

Sociological interest in careers and career patterns is long-standing. A number of early papers traced careers in particular occupations (Hall 1948; Becker 1952), while others considered patterns of careers traversing occupational sectors (Form and Miller 1949). The 1950s and 1960s brought continued attention to individual occupations (e.g., Reissman 1956; Lortie 1959; Smigel 1964), often emphasizing the relatively orderly career progressions among professionals. There also appeared papers on organizational careers (Vroom and MacCrimmon 1968), mirroring the broader fascination with the fate of professionals in organizations.

But by the 1960s studies of intersectoral mobility within careers had merged into the broader mobility literature, which usually dismembered careers into individual transitions (for a review see Boudon [1973]). The 1960s also brought the status attainment paradigm, which collapsed the past career into two or three linear variables predicting current occupational status (but see the debate between Featherman [1971] and Kelley [1973]). As a result, the career as an actual sequence of events became almost invisible to sociological methodologies by the 1970s. On the empirical side, moreover, the orderly careers apparent in earlier studies of professionals proved a mirage; as early as 1961 Wilensky had shown that a large portion, if not a majority, of the work force followed what he termed disorderly careers. Evans and Laumann (1983) were later to show the same for the professionals themselves.

The careers literature moved in new directions in the 1970s. By showing the dependence of career on organizational structure, White's (1970) brilliant analysis challenged fundamental presuppositions of the attainment school. His book was followed by a large literature on formal models of organizational mobility (Stewman 1975; March and March 1977; Rosenbaum 1979*a,* 1979*b;* Konda and Stewman 1980; Stewman and Konda 1983; Skvoretz 1984*a;* Stewman 1986). Simultaneously, practitioners of the status attainment paradigm themselves developed a new emphasis on structure. There resulted a literature on sectoral attainment and career patterns, applying the methodological strategies of the earlier period to the careers of structurally segregated groups (e.g., Rosenfeld 1979; Grandjean 1981; D'Amico 1985). The sectoral approach naturally turned to the transition models of the mobility literature (e.g., D'Amico and Brown 1982; Tolbert 1982), and the focus on interpretation of transitions in turn led to the log-linear version of standard methods (e.g., Gaertner 1980; Wanner and Lewis 1983). Partial-likelihood models have also been used to model job-change rates within sectors (Carroll and Mayer 1986).

This new work did little, however, to resurrect the connection between the successive moves constituting a career. Rosenbaum (1979*a*) did study two-step influences, and Kalleberg and Hudis (1979) have considered whether certain types of "career sequences" (interfirm movement, interoccupation movement, or both) had consequences for further attainment. But while analysts have produced models considering various steps in attainment (e.g., Spilerman 1977; Rosenfeld 1980), they have generally not used information on prior sequence. Although event-history methods enable the use of sequence information, the insuperable difficulties of categorizing prior sequence generally prevented its use as an independent variable. The few attempts at studying whole careers have been strongly theoretical (e.g., Sorensen 1974; Skvoretz 1984*b*).

In many ways, the disappearance of fully linked careers as dependent and independent variables in sociological studies reflected substantial theoretical developments. The central debate of the recent careers literature has pitted human capital theory against the structural determinisms of job vacancies, organizational demographics, and sectoral labor markets. If structuralists were correct that careers took directions set by structural conditions—directions that would appear somewhat random from the standpoint of the individual career—then perhaps the linked career was indeed a mirage.[2]

There are, however, strong substantive reasons for treating the linked

[2] The ideal-typical career as envisioned in the literature advances freely through a set of stages to some goal. In fact, patterned careers could emerge as easily through structural constraints (i.e., typical "failed" careers) as through the open, independent mobility originally envisioned in the career literature.

sequence of moves in a work history as more than an artifact. In the first place, individuals continually plan and structure their work histories, responding to structural constraints with moves of their own. These moves are based on an ongoing sense of occupational self; hence, what seem to be disorderly careers may in fact be logically structured from the individual's point of view (cf. Wilensky 1961). Second, individuals' plans for the future reflect not only their immediate present but also the actual sequence of their experience in the past. Individuals construct their futures not only on an appraisal of their current chances but also by comparing their entire past histories to well-known models (see, e.g., Levinson 1978; Bertaux 1982). There are thus reasons to believe that individuals themselves try to structure their work histories into careers that they find culturally acceptable, into patterns that they recognize.

The very existence of culturally accepted models for careers raises the question of whether the models describe realities or whether, as Rosenbaum (1979*a*) suggests, the ideal-typical career is in fact the rarest of all. Perhaps only job sequences starting in certain sectors can achieve the characteristic linked sequences we usually call careers. Or perhaps, as Wilensky (1961) seemed to indicate, many people lack these linked sequences in order that a few may have them. Such empirical questions all assume an answer to the preliminary question of what typical careers actually look like. They add to the substantive arguments urging us to address the issue of overall career patterns directly.

But the disappearance of overall career patterns from the literature has reflected not only theoretical but also methodological issues. There has been no standard method for analyzing whole sequences of events, while methods for analyzing attainment or transition at a particular point are widespread. Yet the question of patterns among whole sequences arises not only in the careers literature but also in a number of other areas. Are there typical family life cycles? (See Glick 1976; Spanier, Sauer, and Larzelere 1979.) Are there characteristic sequences of professionalization (Wilensky 1964)? Are there characteristic patterns to the unfolding of revolutions (Skocpol 1979)? Are there characteristic "criminal careers" (Blumstein and Cohen 1987)? These questions closely resemble the question of whether individuals really have patterned work careers. Each involves a data set made up of individual sequences of events and asks whether there is a standard pattern or patterns among them. A method that could answer the careers question would work with these questions as well and hence be doubly important. Let us then consider the methods available for answering such "typical sequence" questions.[3]

[3] The notion of ideal-typical sequences for the analysis of historical processes was originally proposed by Weber (1949, pp. 89–112).

II. SEQUENCE METHODOLOGIES

There are a variety of methodological approaches to the central sequence questions of (1) whether there are common patterns among a set of sequences (the pattern question) and (2) if so, how they are produced (the generation question). In practice, the pattern question has proved so problematic that the most common strategy with sequences has been to answer the generation question with an estimated model and then to see whether the model generates sequences whose aggregate properties (not including sequence patterns) resemble those of the original data. (Such properties might include sequence length distributions and average numbers of appearances by particular sequence elements.) The methods we present here, however, work with patterns directly. But before considering these and other direct approaches to the pattern question, we shall briefly consider the relation between the two kinds of questions. That consideration will help clarify the relation of the methods here discussed to more familiar techniques.[4]

A substantial methodological school believes that pattern questions about sequences are uninteresting or misguided. On this argument, sequences like careers are a surface reality generated by an underlying probabilistic process. At each instant the likelihood of an event's occurring to any particular case is a function of exogenous variables and perhaps of time passed since some prior event. (Generally these methods are concerned with the waiting time to one particular kind of event, rather than with the temporal relation of several kinds of events, although so-called competing-risks models are an exception.) In such an approach, questions about typical sequences are false questions, involving not true regularities in the underlying mechanism that produces events, but only surface regularities apparent in the underlying mechanism's results. The sequence is simply the list of events the stochastic process happens to have produced over many time periods.

[4] Abbott (1983) has reviewed sequence literatures up to the early 1980s. Since that time they have developed in a number of directions. Sequential models are used in economics to analyze potential equilibria and optimal strategies in discrete-step competitions (e.g., Dewatripont 1987; Hopp 1987), as well as to analyze the direct dependence of outcomes on path (David 1985). In psychology, there is extensive work on cognitive ability to generate and model sequences (Oppenheimer and Van der Lee 1983; Treisman and Williams 1984; Oppenheimer and Groot 1985; Wyer et al. 1985), on sequences in cognitive development (Campbell and Richie 1983) and on decision sequences in small groups (Poole 1983). An interdisciplinary literature continues the focus on conversational sequences, with some attempts at quantitative analysis (e.g., McClure 1983; Vuchinich 1984). The artificial intelligence community has pursued sequence-generating algorithms (see, e.g., Dietterich and Michalski 1985). Finally, Abell (1984, 1985, 1987) is developing a direct approach to sequence resemblance based on formal homology.

There are, however, reasons for taking the position that typical-sequence questions are of interest. First, and most important, we are often interested in several different types of events—different kinds of jobs, different policy choices—and their sequence. Methods based on the stochastic conceptualization are extremely cumbersome with such problems. Second, some temporal sequences do not, in fact, arise from stochastic processes but are set up as wholes, as scripts. Examples are the sequences of pretrial procedures in different courts, of initiation rites in secret societies, and of types of training in the various professions. About such sequences stochastic models can produce no answers. And moreover the obvious first question about these sequences is whether the (nonstochastic) sequences are similar in the various cases—the pattern question.[5] Even where an underlying stochastic process obviously does occur, as in careers, there are reasons for wondering about typical sequence, as we have noted above. People judge their own careers—and guide their future decisions—partly by comparing their own careers to culturally accepted models. We have no way of knowing either what those models are or whether they are realistic other than by asking sequence pattern questions. There is much work about the accuracy of people's appraisals of their career *chances* (e.g., Rosenbaum 1984, chap. 1); those chances are, of course, what the stochastic methods can tell us. But about people's images of typical careers, or the way their experience bears those images out, we can know little without asking the question directly. A final reason for studying pattern questions directly is that their answers may be useful in stochastic models themselves. To the extent that stochastic models themselves consider effects of the past order of events on present chances, they can make good use of a method for categorizing those past orders.

There are thus a number of reasons for examining sequence pattern questions even if we accept the notion that many important social sequences are stochastically generated. We shall therefore now turn to a review of methods for examining those pattern questions.

One version of the typical-sequence problem has had working answers since the early 1950s: the case where each sequence element can be observed at most once. This situation seldom arises in the careers case; if we observe individuals once a year, most individuals will repeat jobs from year to year. A better example comes from the professions literature, where a classic problem has been whether or not there is a sequence of

[5] These examples all involve multiple events, but one can conceive of single-event durations to which the stochastic model would not apply, like the waiting time to establish citizenship in various states; such durations are of course usually treated as simple nondurational variables.

professionalization, a typical order in which occupations acquire an association, licensing, examinations, schools, and the other accoutrements of professionalism (Wilensky 1964). Another common version is the order of events in the young-adult life course: college, first full-time job, marriage, first child. The general problem in all these applications is to find resemblances among permutational data.

With short permutational sequences, enumeration is a possibility. Thus Hogan (1978) analyzed life-course data using a simple listing of the six order possibilities for his three events. For longer sequences, enumeration rapidly becomes hopeless. One may then use a variety of multilinear permutation statistics; the Spearman rank correlation coefficient, for example, may be viewed as a distance measure between two such permutations and may be generalized in a variety of ways (for a review, see Hubert [1979]).

Multidimensional scaling may also be used to study nonrepeating sequences, an approach long customary in archaeology. (For a comprehensive but dated review of this literature, see Hodson, Kendall, and Tautu [1971].) Archaeologists have often been interested in "seriating" (i.e., ordering) different types of manufacture from measures of resemblance between types—either physical proximities in sites or indicators of stylistic resemblance. Ordering the types of manufacture is simply a one-dimensional scaling problem, although some approaches (e.g., the "horseshoe" method) use two-dimensional algorithms (see also Hubert 1976). With temporal sequence data, seriation techniques can be simply adapted by using various measures of elapsed time to capture the mean "distances" between pairs of events across all sequences possessing those events. One can then order the events overall by submitting these distance data to standard scaling algorithms. Seriation techniques do not directly measure resemblance between two sequences, but rather produce an archetypical or best sequence of the various events. They are not bothered by missing data (as are permutation statistics) but have the usual drawbacks of scaling: dependence on choice of distance measures, volatility with data change, and unfamiliar (to sociologists) measures of goodness of fit. On the other hand, they provide viable answers concerning typical patterns in unique event sequences.

Sequences of repeating events have proved more difficult, for enumeration is impossible and scaling irrelevant. As a result, the general strategy has been to hypothesize a stochastic generating model, to estimate its parameters from data, and then to compare its aggregate properties to those of the sequences observed. By far the most common such generating model is the discrete-state, first-order Markov process. The possible events in the sequence define the states of the process, a start vector describes the potential starting points, and, if necessary, an absorption state sets limits to sequence length. Markov models can be easily estimated from sequence data; the sequences are broken into one-step transitions, and the probabilities of moving from one state to another are calculated directly from these transitions. With suitable algebra, the resultant transition matrices can be directly examined for aggregate properties like expected number of transitions through given states, estimated length and termination state (for absorbing chains), and so on. These estimates can then be compared to observed values to test goodness of fit. Markovian sequence models have several advantages. They reduce the welter of potential sequences to a relatively simple set of individual transitions. And when the particular transitions are treated individually the models allow the introduction of differential exogenous forces, as in event-history methods.

But there are several difficulties. First and foremost, the sequence-generating process may not actually be Markovian (i.e., we may think it has a longer history than the immediate past). For most career models, for example, Markovian assumptions are obviously untenable. If a professor resigns on a whim to become a taxi driver, we know that two or three iterations down the road he or she does not have the same probability of remaining a taxi driver as does one who had been a taxi driver most of his or her past career. Yet the assumption of similar probabilities for all current members of a given state (e.g., taxi drivers) is the foundation of any Markovian analysis. To address this problem by directly estimating deeper dependence (with *n*th order processes) multiplies the size of the state space by itself for each additional time period analyzed. Second, and even more important in the long run, Markov and other stochastic models do not answer the question of whether there is a typical sequence or typical sequences. They do produce a generating model, but we test that model using aggregate sequence properties, not direct measures of sequence resemblance. Finally, Markovian and similar methods are clearly useless in assessing similarity among sequences whose steps were all created at once, as, for example, in the sequential patterns of different versions of a ritual like the Mass (Dix 1945, p. 432).[6]

[6] For reviews of the extensive literature on Markov models of mobility, see Boudon (1973), Stewman (1975, 1976), and Singer and Spilerman (1976). More recently, event-history methods, which test stochastic models of individual, Markovian transitions given certain exogenous forces, have achieved prominence (see Hannan and Tuma 1984). The issue of whether careers "really are" Markovian requires an extended treatment impossible here. Certainly enough doubt exists that we feel justified in proceeding with a method that makes no stochastic assumptions.

There are a number of more specialized techniques for studying sequences, based on more specialized stochastic models. Some, like ARIMA methods, are not really designed to answer sequence pattern questions, but address them implicitly. But even in these a central difficulty remains. For we cannot really estimate the effects of past sequences of events—a central concern of ARIMA models—without an effective means to categorize those past sequences. Even the direct testing of the Markov model—in terms of actual resemblance between generated and observed sequences—awaits a technique for sequence resemblance and categorization. In short, not only would such a technique directly address the issue of finding typical sequences, it would also permit more direct testing of various generating models and facilitate the use of past sequence as an independent variable within such models.

III. OPTIMAL MATCHING TECHNIQUES

The problem of directly measuring sequence resemblance can be solved by a simple dynamic programming technique called optimal matching (sometimes optimal alignment) that is widely known in natural science. Optimal matching measures sequence resemblance when sequences consist of strings of well-defined elements (which may or may not repeat) that are drawn from a relatively small total set of elements; DNA sequences are an obvious example. Optimal matching algorithms do not directly answer questions about sequence pattern; rather, they generate interval-level measures of resemblance between sequences. These measures, taken over a sequence data set, are then input to clustering, scaling, or grouping algorithms, which in turn generate information on typical patterns of sequences. The standard reference on optimal matching is Sankoff and Kruskal (1983).[7]

Suppose that we have data on careers within an organization that has 10 hierarchical levels of jobs. Consider the following career sequences, each listing the level an individual has achieved in successive quinquennia after initial employment:

[7] Optimal matching methods operate by "dynamic programming," a class of iterative maximization techniques operating on stepwise processes. They have been widely used to analyze the resemblance of DNA molecules and to help construct trees of descent among them. The algorithms are also used for computer file comparison and, with the appropriate adjustments, for such continuous-time problems as speech recognition through the matching of sonograms. For detailed references, see the Appendix. For an example, see Abbott and Forrest (1986).

	Years							
	5	10	15	20	25	30	35	40
Sequence:								
A	1	1	2	3	4	5	6	
B	1	2	3	4	5	6	7	8
C	1	1	4	5	7	8	7	7
D	4	5	6	7	7	7	7	7
E	1	2	3	4	5			

These sequences mostly embody upward mobility through the hierarchical levels, each moving in a slightly different way. Sequence C is unusual in skipping levels 2 and 3 and in the later demotion. Sequence D is unusual in starting in the middle of the hierarchy. (Such patterns are obvious on inspection of five sequences but would be harder to find among 500.)

The optimal matching algorithm measures distance between such sequences in terms of the insertions, deletions, and substitutions required to transform one sequence into another. Sequence A can be transformed into sequence B by the deletion of a 1 at the beginning and the insertion of a 7 and an 8 at the end. Sequence A can be transformed into sequence E by the deletion of a 1 at the beginning and of a 6 at the end. We can thus loosely think of A as "closer" to E than to B, since it takes only two actions—the two deletions—to make the change, rather than the deletion and two insertions it takes to make A into B. (This measure of distance is called Levenshtein distance after its Russian inventor. See Day 1984; Sankoff and Kruskal 1983, pp. 18 ff.) It can easily be seen that the numbers of transformations required to turn these sequences into each other are as follows:

	A	B	C	D
B	3			
C	6	5		
D	8	7	5	
E	2	3	6	8

Note that, while in principle there are many different ways to turn one sequence into another, the numbers here are the *minimum* number of transformations required. Even so, there are often several different ways to make the change. Consider sequences A and C. Obviously any pairing will take advantage of the first two elements, which are the same in each. But from there there are several possibilities. We can substitute 4 for 2, 5 for 3, 7 for 4, 8 for 5, and 7 for 6, and then insert a 7 at the end, at a cost of 6 changes total. Or we can insert the extra element as the 4 after the identical opening pair, and then make five unmatched substitutions to the end. Or we can delete the 2 and the 3 (cost of 2), match up the 4 and 5 (for no cost), substitute 7 for 6, and then insert 8, 7, and 7. This can be a little more easily seen in the illustration below, which shows these "alignments" using the letter φ (for null) to hold a place in one sequence when we are inserting in the other. Each vertical pairing in the alignment embodies a transformation, all of which have unit "cost" except those pairing identical elements.

1	1	4	5	7	8	7	7
1	1	2	3	4	5	6	φ

1	1	4	5	7	8	7	7
1	1	φ	2	3	4	5	6

1	1	φ	φ	4	5	7	8	7	7
1	1	2	3	4	5	6	φ	φ	φ

In every case, the cost of the total transformation is 6.[8] It is clear that, as the sequences become more lengthy and complicated, inspection will not always produce the correct answer.

There are three further complexities. First, it is clear that the length of the sequences influences the number of transformations required. And in a sequence of three steps, one substitution is more important than in a sequence of 20. We thus standardize by dividing the transformation distance by the length of the longer sequence. (There are other strategies, but that one is followed here.) This gives the following distance table.

	A	B	C	D
B	.375			
C	.750	.625		
D	1.000	.875	.625	
E	.286	.375	.750	1.000

This table produces a reasonably intuitive set of distance measures between the sequences. Sequences A and E are quite close, A and B as well as B and E are a little less close, and A and E are both far from D. Looking back at the original data, we can see that A and E are people who are working their way up from the bottom and have not yet succeeded, while D has worked from the middle to the top and stayed there. Optimal matching finds this distinction effectively.

[8] There are in fact four more ways to do this, one for each placement of the null element between its locations in the top two alignments. Note that the Levenshtein distance measure is bounded, since the worst-case transformation between a pair of sequences involves only as many substitutions as the length of the shorter sequence and as many insertions as the difference between the two sequence lengths.

A more important complexity is that not all substitutions really "cost" the same. The difference between senior executive and line worker is greater than the difference between first-level supervisor and line worker. So we must differentiate the costs of substitution. This is a theoretical task, and a central one in any application, as we shall show below. However, for the purposes of the present example, a simple assumption might be that all costs of substitution are proportional to the absolute value of the difference between levels; big jumps cost more. This assumption of course will require that we recalculate the intersequence differences.

A final issue is that insertion and deletion costs may themselves vary; because being president is unusual, perhaps inserting it should cost more than inserting another position. However, these costs are in some sense a function of what the sequence already looks like; being president would be unusual for some careers but not for others. Because of this uncertainty, many applications make insertion and deletion cost the same in all cases and set that cost to a value equaling or slightly exceeding the highest cost of substitution. We follow that practice here.

These decisions about costs, of course, can be changed once one sees how the algorithms are working. Their purpose, after all, is to help us categorize the sequences, and we judge them by whether they produce useful categorizations of careers. It is, of course, likely that the most useful categorizations will arise when we use the most substantively sound set of costs, as we shall see.

Allowing the costs of substitutions to fluctuate makes the job of finding the minimum cost transformation considerably more complicated. Inspection is no longer effective. The problem is solved by a dynamic programming algorithm discussed in the Appendix. The algorithm produces, for any pair of sequences and fixed set of substitution costs, a minimum cost for transforming one sequence into the other. That cost is a distance measure that we can use to classify the sequences.

IV. A SUBSTANTIVE EXAMPLE: CAREERS OF GERMAN MUSICIANS, 1660–1800

To illustrate the power of optimal matching with a substantial data set, we report here on career data drawn from a larger project on the labor market for musicians in Germany during the Baroque and Classical eras, from roughly 1660 to 1810. These careers resemble modern careers in many ways. They not only show various progressions but also traverse organizational hierarchies and rise to status peaks. They contain extensive sectoral mobility—between musical and nonmusical work and between various sectors within music itself. And they are constrained by structure, since few positions are available at the pinnacles of the various musical establishments. In short, these centuries-old careers follow the complex patterns familiar from the highly differentiated and structured labor markets we often believe unique to the present century. While their content may be unfamiliar, the issues they raise are familiar indeed.

To help conceptualize these careers, we shall briefly sketch the job system containing them. Although music pervaded 18th-century German life generally, music in the towns enjoyed some freedom, while music elsewhere was shaped largely by the local princes. We shall therefore review the musical labor market under the headings of court and town, closing with a mention of other musical venues.[9]

Concentration of wealth made the courts a central support of musical patronage. Nonetheless, the musical establishment and the professional musicians occupied precarious positions. Considered servants, court musicians often performed nonmusical duties as well. An unmusical heir or a sudden war might empty the music budget. Financial restrictions often delayed salaries for years at a time. Patrons were many, but permanent, secure positions were few.

The main musical ensemble at a substantial court was the *Kapelle;* its director, the *Kapellmeister,* oversaw the entire musical establishment. In all but the largest courts, the kapellmeister was at once composer, conductor, and performer. At the great courts he was assisted by subordinate musical administrators, customarily a musical director, a composer, and a *Konzertmeister.* There were also a variety of music teachers and, in the larger courts, perhaps a separate opera establishment. Most courts, of whatever size, also employed an organist for religious services. Large courts generally had long musical traditions and employed many resident full-time musicians, sometimes up to 100 or more. Smaller establishments were more precarious. But in both cases the patron and the patron's finances determined the musicians' lives. Demand for music might suddenly appear and just as suddenly vanish.

Although the musical work in the towns was differently structured, there too music was a universal practice. Yet though music was of great civic concern, town musicians often needed considerable outside income, both musical and nonmusical, to supplement meager salaries in the civic, church, or school musical establishments.

[9] This section rests on an extensive review of secondary literature as well as on generalizations from careers we have coded. General sources are Raynor (1972), Hogwood (1977), and various essays in Salmen (1983), in *Die Musik in Geschichte und Gegenwart* (Blume 1949–79), and in *The New Grove Dictionary of Music and Musicians* (Sadie 1980). More extensive references will be found in Hrycak (1988).

In major towns, the main musical establishment was controlled by a town music director, who oversaw the music in local churches and schools. Although expected to compose and produce new works with frightening rapidity, the town music director might also have various nonmusical duties, often as cantor (headmaster) of the principal town school. The position's onerous duties and comparatively small salary were offset by its stability. The town music director was assured of being paid and, unlike the court music director, commonly had the autonomous right to resign his post.

Town churches permanently employed at least an organist and a cantor, and large churches also employed a music director (usually called kapellmeister). The cantor conducted the local school choirs during the choral parts of the church service, while the organist composed and performed instrumental works independent of the choir. Like other civic musicians, organists and cantors supplemented their salaries in various ways, some by teaching music, some by acting as church secretaries or legal advisers, some by renting out pews with prime musical locations.

The final civic institutions employing musicians were the universities and other schools. University musical life required at least two professional musicians, an organist for the university church and a music director for supervision of musical activities. In addition, there were student musical organizations (*collegia musica*) and student resident houses (*Bursen*) that employed professional (but often part-time) music directors to direct their amateur performances.

Town musicians often crossed institutional boundaries. Church, school, and town ensembles depended heavily on one another for different services. But the various institutional boundaries did make the town musicians a more independent and less hierarchical group than the court musicians, a factor that complemented their greater job security.

Two areas of music lay outside the official lines of court and town: opera on the one hand and music for private citizens on the other. Opera was a relatively new area of employment. Although court operas dated from the turn of the 18th century, noncourt opera companies were few and impoverished. Supplementing these were numerous itinerant companies, each employing a few musicians who acted, sang, and accompanied, and each directed by an opera manager who was at once entrepreneur, employer, and musician. Public opera was the first major form of free-lance work in music and the most dependent on immediate public response.

Finally, a substantial group of musicians provided music for hire to private individuals, generally well-to-do citizens in major towns. Some of these performers were musicians' guild members working on a per diem basis. Others were not guild members, often itinerants without citizenship. Few guild musicians, and indeed few of the lesser town musicians, appear in our data.

The majority of German musicians in the 18th century thus worked in one of two spheres, the court or the town. While the civic sphere had neither the high salaries and status of the court nor the excitement of the opera, many musicians chose civic employment for its stability. Operas and courts were transient, while the school, the university, the church, and the town hall were not. Moreover, town musicians were often appointed for life, while other musicians depended on the constant approbation of a demanding, supercilious, and not occasionally impecunious patron or public. A town musician was his employers' social equal; his extra duties, while onerous, lessened his risks of musical failure. In short, the court (and to some extent the opera) promised fame and fortune to a select few for a varying period. The town offered a sure foundation, usually for life.

The career patterns we expect in such a labor market ought to reflect the incentives and opportunities just outlined. At a minimum, we expect a small number of stable court careers and perhaps a larger number of stable town careers. The difference in security predicts a drift from court to town in many careers. It is an open question where the major administrators—kapellmeisters and others—were recruited, whether they came from lower ranks within their own institutions or from some one special source inside or outside. But the importance of the positions and the relative bureaucratization of most courts lead us to expect fixed ladders to these administrative jobs. Beyond that, we have little idea whether there were any fixed career ladders leading through this market. Our chief aim in applying optimal matching techniques to it is thus to find out what categories of careers really do exist within it.

V. THE CHARACTERISTIC CAREERS OF MUSICIANS IN GERMANY

Applying optimal matching takes several steps. We begin with a description of the data and the major coding decisions. (In effect, this is the task of conceptualizing the career sequences, analogous to conceptualizing the variables in standard methods.) We then set the parameters for the optimal matching algorithms, in particular the substitution costs between various types of job. Next we report and analyze the results, using a number of strategies for testing the effectiveness of our derived "model careers." We close the section with a further analysis that employs different models for career time, an analysis reinforcing our earlier results.

Data

Data on the careers of musicians in this labor market were gathered from two musicological dictionaries, *Die Musik in Geschichte und Gegenwart* (Blume 1949–79) and *The New Grove Dictionary of Music and Musicians* (Sadie 1980). The job history of every musician listed in those sources and active in Germany between 1650 and 1810 was included. For each incumbency, we recorded the name of the individual, the position, the ensemble, the employing institution, and the town as well as the dates of arrival and departure. Inclusion in the present study required that musicians' careers be completely known; we had 595 such careers.[10]

The career of an individual started with his first posteducational position in the profession. It ended when the individual permanently left the career system, generally through death or retirement. Dates of moves were coded in years; there was seldom more detailed information. The individual records thus consisted of moves from one job—geographic location, employer, position, sphere—to another, identified by the date of move. Identification (ID) numbers linked the records into full careers.

The positions and spheres of employment for musicians are listed in table 1. Of the 135 ($= 15 \times 9$) position/sphere combinations, only 34 contained more than nine incumbencies among our 595 musicians. We thus chose to regard those 34 as our central "alphabet" of jobs for these careers and to treat the remaining 101 position/sphere combinations as an "other" category. The 35 jobs (34 identified jobs plus the other category) held by these musicians, the numbers who held them, and the percentage they represented of all incumbencies are shown in table 1. (We hereafter refer to these position/sphere combinations as "jobs.")

Another coding decision involved the use of geographic mobility information. With 304 geographical locations, 573 employers, and 35 jobs in the data set, we obviously could not create substitution costs for individually identified jobs. We therefore were forced to conflate "internal" mobility—geographic or employer mobility without job category change—and nonmobility. A continuous listing of CRTKPM means not that an individual stayed with a particular employer at a particular geographic location, but rather that he stayed within the court kapellmeister job type if

[10] Our period opens with the reestablishment of princely patronage and town institutions in the wake of the Thirty Years' War and closes with the Napoleonic cataclysm, which coincided with (and partly caused) the expansion of a new, middle-class market for music. We defined "Germany" linguistically, including, e.g., only the German-speaking lands of the Austrian Hapsburgs and excluding Germans working abroad. Information supplementing the main sources came from court records, published reminiscences, and obituaries, as well as from period reference works.

TABLE 1

POSITIONS, SPHERES, AND JOBS

Title	Final Jobs			
	Label	Number	Percentage	Cumulative
Court instrumentalist	CRTINS	353	16.8	16.8
Court kapellmeister	CRTKAP	280	13.3	30.1
Church organist	CHUORG	233	11.1	41.1
Court konzertmeister	CRTKZM	107	5.1	46.2
Court organist	CRTORG	96	4.6	50.8
Court vocalist	CRTVOC	85	4.0	54.8
Court administrator	CRTADM	74	3.5	58.3
Church cantor	CHUCAN	68	3.2	61.5
Court composer	CRTCOM	55	2.6	64.2
Church kapellmeister	CHUKPM	48	2.3	66.4
Opera administrator	OPEADM	47	2.2	68.7
School teacher	SCHTCH	44	2.1	70.8
Church administrator	CHUADM	41	1.9	72.7
Town music director	TWNMDR	36	1.7	74.4
Court opera administrator	COPADM	33	1.6	76.0
Court nonmusician	CRTNON	33	1.6	77.5
Church music teacher	CHUMTC	31	1.5	79.0
Court accompanist	CRTACC	30	1.4	80.4
Opera kapellmeister	OPEKAP	27	1.3	81.7
Court opera nonmusician	COPNON	26	1.2	83.0
Opera instrumentalist	OPEINS	21	1.0	84.0
Court other/outside	CRTOTH	21	1.0	84.9
Town cantor	TWNCAN	20	.9	85.9
Court music teacher	CRTMTC	20	.9	86.8
Town band administrator	TWNBAN	19	.9	87.7
Church vocalist	CHUVOC	18	.9	88.6
Court opera music teacher	COPMTC	16	.8	89.4
Church nonmusician	CHUNON	16	.8	90.1
Town instrumentalist	TWNINS	15	.7	90.8
Court opera kapellmeister	COPKPM	14	.7	91.5
Town organist	TWNORG	12	.6	92.1
Opera vocalist	OPEVOC	10	.5	92.5
Opera composer	OPECOM	10	.5	93.0
Court opera vocalist	COPVOC	10	.5	93.5
Everything else not in this list	OTHER	147	7.0	100.0

NOTE.—Positions: vocalist, instrumentalist, composer, accompanist, organist, kapellmeister, konzertmeister, cantor, music director, administrator, teacher, nonmusician, itinerant, amalgamation, outside. Spheres: court, court opera/harmonie/special group, aristocrat without court, academy/school, town general, town special group, church, theater/opera, other. There are 101 possible positions in the OTHER category.

he did move. We are therefore viewing careers as successions of types of jobs, rather than of particular jobs themselves.[11]

A final decision involved the amalgamation of jobs. A startlingly large fraction (about half) of these musicians held two jobs concurrently at some point in their careers. Such amalgamation might involve one extra job for a year or an entire secondary career line. The handling of secondary career lines presents obvious problems for any method treating careers as sequences, whether optimal matching, event history, or direct Markovian. In other work, we are addressing this problem by classifying actual job pairings into types that can be introduced as elements in a single-line career. But the details of that discussion would take us well beyond the bounds of the present article. We have therefore used only those careers not involving amalgamation in the optimal matching analysis ($N = 279$).[12]

Each musician's career was thus expressed as a sequence of the 35 jobs, using one year as the basic time interval. In the examples that follow we avoid long lists of identical positions by listing consecutive years spent in a position before the position's name; 38CRTINS means 38 years as a court instrumentalist. The examples include an eminently successful career in court composition, a variegated career leading to a prominent town position, and a short and relatively unsuccessful court career. We include two examples with amalgamations to illustrate the potential complexity of amalgamated careers.

[11] This problem of internal mobility is not specific to optimal matching studies. It is familiar from the conflation of "stayers" and "movers within category" on the main diagonal of mobility matrices. As probability theory predicts, we found the holding of equivalent jobs for different successive employers to be most common among the most common jobs. The top six jobs (CRTINS, CRTKPM, CHUORG, CRTKZM, CRTORG, and CRTVOC) provide 75% of these "consolidations" and follow the same order as generators of consolidations as they do in jobs. There is thus little evidence of biasing through particular jobs' being more or less likely to conceal such internal mobility. A substantial fraction (about 40%) of the careers had at least one consolidation.

[12] That J. S. Bach and Telemann combined several jobs suggests that leading musicians might more likely be amalgamators. But counterexamples like C. P. E. Bach and Haydn are common. We have calculated the distributions of types of jobs in amalgamated and nonamalgamated careers and find, in accord with our earlier substantive discussion of the relatively formal and bureaucratic character of court employment, that the commoner court jobs (instrumentalist, kapellmeister, konzertmeister, vocalist) are slightly overrepresented in the unamalgamated careers. (These four constitute about 50% of all jobs in such careers as opposed to 40% overall.) Although this does not seem a substantial imbalance, we will of course reanalyze the data once a full categorization of amalgamations is derived.

A. Salieri:					
Main	50CRTCOM				
Amalgamated	16COPADM	36CRTKPM			
J. S. Bach:					
Main	1CRTVOC	5CHUORG	10CRTORG	6CRTKPM	27TWNMDR
Amalgamated	1CRTNON		4CRTKZM		27TWNCAN
					7COPKAP
					8TWNBAN
W. Mozart:					
Main	6CRTKZM	2OTHER	2CRTORG	5OTHER	4CRTCOM

Optimal Matching Parameters: Substitution Costs

After defining the sequence elements, the next task in optimal matching is to set the costs of substitution between those sequence elements, in this case the different jobs. As we shall show, the setting of substitution costs involves central theoretical questions.

In our case, the initial basis for setting the costs is clearly the relative importance of changing sphere on the one hand or position on the other. Some job moves involve a change of sphere, some a change of position. There is little conceptual basis for arguing that a move from court to town is more or less drastic than a move between positions within one or the other world. The data show numerous moves across both sphere and position boundaries. It is, however, clear that a change in both sphere and position is a more drastic difference than a change in only one or the other. We thus assigned substitution costs of .75 between any two jobs involving only a position change or a sphere change and 1.0 between jobs involving a change in both.

However, it is clear that some pairs of jobs were fairly closely connected by mobility; they seemed often to lie on the same career lines.[13] We

[13] Using such mobility information to estimate substitution costs raises a serious question. By using it we translate diachronic closeness into synchronic closeness, which we then use to estimate diachronic career resemblance. In one sense this procedure "programs in" the common career patterns ahead of time; jobs linked by mobility will have low substitution costs, and careers involving them will then be linked closely by the matching algorithms. However, there are strong arguments for using the mobility information. In the first place, the general categories already used to establish costs—"sphere" and "position"—are themselves to some extent mobility categories. The substantive closeness of church organist and church cantor lies not only in the common sphere of activity, with all that means about financial security, community prestige, and so on, but also in the networks of acquaintance that facilitated motion between one position and the other. Moreover, our coding scheme already uses mobility information. For by consolidating comparable jobs with different employers, it effectively regards transition between those jobs as costless. Finally and more generally, if we wish to find characteristic career patterns and already know that certain moves are more common than others, then surely we wish the matching algorithms to take ad-

have therefore added mobility information to our measures of job resemblance. This involved (*a*) producing a "distance" matrix based on mobility information and (*b*) combining that matrix with the position/sphere dissimilarity matrix. To do the first of these tasks, we began by creating a matrix of transitions between the 35 jobs, classifying in it all moves in all careers. Since the substitution matrix (into which this information would eventually go) must be symmetric, we then symmetrized this matrix by adding corresponding elements (i.e., the i,jth and the j,ith for all i and j) and replacing both the i,jth and j,ith elements with this sum. Since the last row and column of this matrix (the "other" category) actually represented 101 different types of position/sphere combinations, we divided the elements of this row (column) by 101 to weight them proportionately to the actual mobility involved. Finally, since transition figures rise with mobility closeness rather than with distance (i.e., cost), we turned the matrix into a dissimilarity matrix by subtracting each element from a constant. To make a composite measure of dissimilarity, we then combined this transformed transition matrix with the position/sphere matrix through a linear combination of corresponding elements in the two matrices. We chose coefficients for the combination so that *maximal* mobility "similarity" caused the same reduction in substitution cost as did sphere or positional similarity.[14]

Setting substitution costs thus involves serious reflection about available information as well as careful consideration of assumptions about unavailable information. Yet, while care is required, experimental work by Abbott and Forrest (1988) indicates that the calculated distances are not substantially affected by even fairly strong perturbation in the substitution costs. Abbott and Forrest had five judges independently code both the sequence elements and the sequences in an identical data set of 40 dance sequences. Judges varied considerably in their definitions of

vantage of that information. The situation is analogous to the use of observed transition information to parameterize models in standard stochastic studies of mobility.

[14] Various means can be used to mix the different types of resemblance of sequence elements into a composite measure of interelement distance. If one has a number of properties whose presence or absence characterizes each element, one may use Jaccard, matching, or other binary distance measures; at a very simple level (two properties) that has been our procedure with sphere and position information. If one has several different resemblance matrices, one can create a linear combination of them; that has been our procedure here in combining the position/sphere and transition measures. In our substitution-cost matrix, 425 of the 595 elements equaled the maximum substitution cost of 1.000. The smallest cost (closest resemblance) was .47 between court instrumentalist and konzertmeister, reflecting the common sphere of activity and the common mobility between the two. We have set the insertion/deletion cost to exceed the highest substitution cost (1.0) by the difference between the highest and next highest substitution cost (insertion/deletion cost = 1.0188).

elements and in their substitution-cost designs. Some saw as many as 81 different figures in the dances while others saw as few as 29. The classifications of these figures into hierarchies (from which the substitution-cost matrices were calculated) were equally diverse. Yet the distance matrices produced by the optimal matching analyses from each judge's sequence coding and substitution design were overwhelmingly similar across all pairings of judges. (Abbott and Forrest [1988] employed both Monte Carlo tests and direct permutation techniques to assess this similarity.) The resemblance persisted in the scalings and clusterings produced from the individual judges' codings of the data; these, too, showed overwhelming similarities. The method thus seems to behave robustly with respect to variation not only in substitution costs, but even in the splitting or lumping of particular types of sequence elements (i.e., in the jobs, in the present case). As is often the case, while care is needed, differences in minor analytic decisions are unlikely to drastically change results.

Results

Given a model of substitution costs we can apply the optimal matching algorithm to the data. For reasons that become clear below we have divided the data into three randomly chosen groups.[15] Our first group contains 94 sequences, of which five are duplicates. For these 94 sequences, the optimal matching algorithms were used to produce an intercareer distance matrix. Since inspection showed an obvious variety of career patterns, our chief aim was to produce a viable classification of careers from this matrix, rather than to find a single typical sequence. It is well known that different clustering algorithms produce different kinds of classifications from complex data (see, e.g., Hartigan 1975). Therefore, we applied several algorithms and created categories that seemed effective across them. We then verified these categories with formal methods.

The detailed procedure was as follows. We applied three different clustering methods to the matrix: single-, complete-, and average-linkage clustering. Then, supplementing the three clusterings with our own substantive knowledge, we classified the careers into 20 different career types. We also allowed an unclassified group for careers that did not seem to follow any clear pattern. Table 2 presents statistics on the groups so

[15] We have divided the data into three groups in order to compare three different models of time later in the article. We do not analyze the entire data set here because we are using one of those models of time and feel, conservatively, that it is not legitimate to use the same data twice. There is no limitation in the software (e.g., data set size) that prevents us.

derived. The career types are loosely grouped into organist careers, which cluster together, and then into careers in the court and church spheres. (There were no clear groups of town or opera careers in these 94 sequences.) Each career type is given a name roughly defining it. The name usually begins with a temporal designation indicating the duration of the career or of its first segment. The segment(s) of the career are then listed. The table includes the number of careers observed in each group and the average distance between careers within group. (The distances are scaled so that the maximal distance between two careers in this data set is 1.0.) The right-hand columns contain an "ideal type" of the career, what we feel is the typical version of that career. Since the groups were small, these were derived by inspection. They could, of course, be derived in a variety of formal ways: by scaling each group individually and choosing the most centrally located career, by creating hypothetical careers that would be centrally located in such a scaling, by choosing the career with the smallest average distance to others, and so on. A typical entry, then, is that for "short-term instrumentalist" careers at court. There were six of these, with an average distance of 0.309 between them; the ideal type was 15CRTINS: that is, 15 years as a court instrumentalist, with no other career service.[16]

We have formally tested these groupings in two ways. First, we have examined the hypothesis that mean distances within groups are substantially less than mean distances between groups. Since distance distributions are ill behaved, we have used a jackknife procedure; the results are reported at the bottom of table 2. It is clear that the derived groups are indeed an effective grouping of the data. A second test for the validity of the groups is based on the ideal-type careers. We have measured the distance from these 20 patterns to the 88 careers actually placed in categories (six careers were unclassified). Eighty-five of these careers are closer to the model proper to their category than to any other ideal type.[17]

[16] In the clustering procedure, we attempted to steer between two poles: on the one hand simply labeling the clusters as they appeared on printouts, on the other hand using prior substantive information to justify arbitrary reinterpretation of the clusters. The validity of our procedure seems supported by the jackknife ratio tests reported below. We chose 20 as a target number of different careers because figures substantially lower lumped careers we wished to distinguish and those substantially higher seemed little improvement on the original data. Such a choice—formally the choice of cluster diameter—is always dictated by substantive considerations. For example, there are, under all three methods, three very general clusters of (1) organists, (2) other court musicians, and (3) other noncourt musicians. In another study, that more general clustering might be the clustering of interest.

[17] The jackknife procedure estimates variances by calculating sample variances on the N samples of size $N - 1$ from which one data point (successively) is deleted. Since distance data are truncated at zero, a logarithmic transformation is performed before calculation.

TABLE 2

CAREER PATTERNS OF GERMAN MUSICIANS: REAL MEASURE

Type of Career	Number in Group	Average Distance	Ideal Type	
Organists:				
Long-term court organists	3	.397	40CRTORG	
Long-term church organists	9	.229	40CHUORG	
Short-term organists	4	.484	15CHUORG	
Court system:				
Kapellmeisters:				
Short-term kapellmeisters	5	.404	10CRTKPM	
Medium-term kapellmeisters	5	.404	20CRTKPM	
Long-term kapellmeisters	8	.356	40CRTKPM	
Vocalists:				
Medium-term vocalist	4	.501	23CRTVOC	
Vocalist/kapellmeister	2	.331	25CRTVOC	16CRTKPM
Instrumentalists/konzertmeisters:				
Short-term instrumentalist	6	.309	15CRTINS	

Medium-term instrumentalist	9	.285	25CRTINS		
Long-term instrumentalist	5	.292	50CRTINS		
Short-term instrumentalist/konzertmeister	5	.325	5CRTINS	30CRTKZM	
Instrumentalist/konzertmeister/administrator	2	.457	8CRTINS	20CRTKZM	23CRTADM
Instrumentalist/other/kapellmeister	3	.491	10CRTINS	10OTHER	10CRTKPM
Other court positions:					
Court composer	3	.614	40CRTCOM		
Opera administrator/court opera administrator	3	.645	15OPEADM	8COPADM	
Church system:					
Church kapellmeister	2	.442	25CHUKPM		
Church cantor	2	.541	30CHUCAN		
Other jobs:					
Short-term other	2	.532	10OTHER		
Long-term other	6	.440	40OTHER		
Unclassified	6	.913	Other		
Average within groups		.372			
Average between groups		.860			
Jackknife ratio estimate		1.495			
t-value		10.31			
df		93			

The ideal types offer the analyst a powerful tool. Like all distance-based techniques, optimal matching methods require dyadic (relational) data; each *pair* of points in a sequence data set generates an observation. Since the number of distances to be calculated rises with the square of the data set size, data sets must be relatively small. But ideal types free us from this constraint. If we use pilot data to establish a number of basic patterns, we can then use those patterns to classify a larger data set with that number of measures per case. By developing thresholds for classification (that is, by setting a diameter within which a sequence must lie to enter a category), we can create a residual category if necessary. The development of ideal sequence types thus offers an escape from limitations inherent in optimal matching as a relational method.[18]

An Elaboration—Time Warping

Optimal matching techniques permit a variety of elaborations. One of these involves varying our assumptions about the meaning of time. Our normal assumption is that one year is equivalent to another year, that

[18] The idea of using ideal types to evade the dyadic data problem arose in a conversation between the senior author and Alfred Blumstein of Carnegie-Mellon University. To our knowledge, such types have not been used in previous applications, which are mostly biological and concern relatively small numbers (dozens) of extremely long (thousands of elements) sequences. Although the use of ideal sequence types is not an absolute necessity until the data set reaches several hundred (and many sequence problems involve data sets within this limit), one may wonder about the information lost when types are used. That information loss can be assessed by a measure like the stress statistic minimized in multidimensional scaling. Every pair of sequences classified by the ideal types has two distances associated with it: the observed distance between the two points and the distance between the ideal types under which they fall. We first create a measure of total disparity by taking the difference between these two distances, squaring it, and summing the squares across all pairs of classified sequences. To consider the ratio of this sum to the actual distances themselves, we then divide this total by the sum of the squares of the actual distances. We then take the square root of the whole expression to return to the original scale. The measure (which is effectively the Kruskal stress measure) has two useful properties. If all sequences are treated as being in one group, the interideal-type distances are all zero, and the measure reduces to unity. In this situation, of course, we have lost all information about sequence differences. If, on the contrary, all sequences are treated as unique groups, the disparities are all zero and the measure itself is zero. Here we lose no information, but of course gain no reduction in data set complexity. For any given set of ideal types, then, values close to zero indicate relatively good fits, with little loss of information. The value of this statistic for the career ideal types just calculated is 0.141, which seems to indicate that the 20 groups retain most of the information in the original 88 sequences. (Of course extensive research would be necessary to set conventions for acceptable values of this measure.) In any case, we feel that ideal types can extend the methods considerably and that the dangers attending their use can be assessed and hence minimized.

temporal measurement is not a problem. But consider two careers in which exactly the same succeeding jobs are held, but in one career for twice as long as in the other. Substantively, we might want to think these careers close; the same processes may be happening but at different rates. Or consider two careers identical for the first 20 years, one of which then ends in death while the other continues the terminal job for another 20 years. Substantively, what matters may be the strong resemblance of the initial sections, not the difference caused by the length of the terminal job.

Our intuitive use of clock time thus contains a hidden measurement assumption that causal processes work at constant rates at all times in all careers. Although this assumption is, of course, common to all measurement systems using clock time, optimal matching allows us ways to relax it. While advanced algorithms do so by changing the meaning of insertion and deletion (see, e.g., Kruskal and Liberman 1983), a simpler method is to transform the data.

In our previous model, we assumed that two careers were substantively different unless they had identical job sequences with identical terms in each job. Careers were thus considered close when job terms varied slightly or when moves were made to similar (low substitution cost) jobs. We now vary these assumptions in two different ways. Our second model considers only the *relative* proportion of jobs to one another within each career; all careers are standardized to a 50-unit length. A short and a long career may thus be identical if they consist of identical jobs in identical order with identical *proportions* of time spent in each job (e.g., the real-time careers 10CRTKPM and 20CRTKPM both become 50CRTKPM). Careers are different when they have different rhythms or different jobs, but not when they differ only in overall pace. The standardized metric thus assumes that causal forces operate at equivalent ratios within careers; a late bloomer will continue to develop slowly (and a prodigy to move quickly) throughout the career. Our third model allows elapsed time to retain some but not all of its power by using a logarithmic transformation of the elapsed period in each job (10CRTKPM becomes 3CRTKPM and 20CRTKPM becomes 4CRTKPM; we add one to take care of stays of one or two years). Here the difference between long and short tenure in a given job makes a difference, but much less than in the real-time metric. In particular, differences are relatively stressed when they involve short periods; the difference between 2 and 4 years on a job becomes more important than the difference between 10 and 7 years or 20 and 25 years. The logarithmic metric thus assumes that causal processes operate at equal rates across careers in the short run (because it makes small differences important) but at varying rates in the long run.

Since we wished to consider these time measures independently, we split the career sequences ($N = 279$) randomly into three groups and transformed them according to the preceding discussion. (The first of these groups is the 94-sequence group already discussed; this is the division referred to in n. 15 and elsewhere above.) The transformations made a number of careers identical that had not been so before, so there were a total of 65 (92 − 27) distinguishable standardized-time careers and 78 (93 − 15) distinguishable log-time careers. These data were then analyzed with the same procedures as reported above. The results are reported in tables 3 and 4.

It is clear in both cases that the categorizations are effective; the jackknife t-values are large. As for the model careers, they correctly classify 84 of 86 classified cases under the standardized metric and 78 of 80 classified cases under the log metric. The stress statistic (see n. 18 above) shows that in both cases the ideal types retain most of the original information.

Many career groups appear across all three metrics. These are the strong candidates for career types. First, there are careers dominated by a single type of job: as court or church organist; as court kapellmeister, instrumentalist, and possibly composer; and as church kapellmeister and cantor. A number of more patterned careers appear in all three metrics (sometimes with varying interludes): court vocalist to kapellmeister, court instrumentalist to kapellmeister, court instrumentalist to konzertmeister, and court instrumentalist to konzertmeister to court administrator. These are the possible career ladders. It is striking that all of them lie within the relatively bureaucratic court system. The entry-level positions of court vocalist and instrumentalist do occasionally lead to substantial advancement, but only within the court itself. There is no hint here of a general drift from the insecure court world to the more secure town and church jobs. The spheres are so strongly separated that no multijob career patterns cross sphere lines. Some careers did of course cross these lines—J. S. Bach's is an example—but they were uncommon. We see, then, that the security of town positions did not really form a systematic attraction for musicians tired of court competition, and that, in fact, court musicians had at least a few ladders to climb. These results strongly revise the loose predictions we derived from the secondary literature.

VI. ADVANTAGES AND DISADVANTAGES OF OPTIMAL MATCHING METHODS

Since this article is chiefly concerned with introducing and illustrating a technique, we shall close with a discussion of strengths, pitfalls, and weaknesses. These determine when optimal matching is the preferred approach to sequential data and when it is not.

The chief strength of optimal matching is its ability to directly measure sequence resemblance. It provides a way of addressing such fundamental questions as whether there is or are common sequential patterns among data. The career types we have found were no doubt loosely recognized by 18th-century German musicians as the few alternative careers possible in their labor market. Markovian analysis cannot give us that list of types; it can only accept or reject Markovian assumptions about careers. Event-history models also cannot develop that list; they can effectively model only one transition at a time.[19]

The patterns derived by optimal matching are amenable to further use either as dependent or as independent variables, depending on whether our further questions take the form of "Why do certain kinds of people end up with certain kinds of careers?" or "Why do certain kinds of past career patterns tend to lead to differing patterns in the future?" In substantive situations in which one may reasonably assume that the sequence patterns developed from one another as wholes, the sequence patterns may be arranged into trees of descent. If, for example, we think certain professions have modeled their development on that of medicine and others on that of law, we may see whether the actual histories of professional development in fact fall in such lines of descent.[20]

This strength gives optimal matching a great breadth of application. One can categorize life-event sequences to see whether certain sequences characteristically lead to stress (as Chalmers [1981] has recommended). One can follow Stinchcombe's (1978, pp. 13–16, 89–97) injunction to develop ideal-typical sequences of historical development. One can study the developments of organizations, of revolutions, of families. Nor does

[19] Of course, it is quite unlikely that musicians would know exactly what typical careers looked like. Rosenbaum (1984, pp. 13–16) notes a long literature on employee misperception of careers, although most of this concerns misperception of opportunities (i.e., chances at a particular point in time) rather than entire careers. Nonetheless, without any means to identify typical careers ourselves, we have no grounds for actually judging employees' perceptions of typical careers.

[20] Optimal matching techniques have no models for prediction in themselves. Categorized sequences can be used as independent variables within any standard technique—e.g., event-history techniques if the concern is the duration of some present event, regression techniques if the concern is some particular outcome. Categorized sequences are appropriate dependent variables under some conditions. Script sequences (temporal sequences set up as units) are legitimate dependent variables, as are stochastic sequences in the case where the determining parameters are presumed fixed from the outset, a condition implicitly argued by some economists studying life-course behavior (see, e.g., Heckman 1978, esp. p. 205; Heckman and MaCurdy 1980; MaCurdy 1981). Otherwise, dates of measurement for independent predictors become a major theoretical problem.

TABLE 3

CAREER PATTERNS OF GERMAN MUSICIANS: STANDARDIZED MEASURE

Type of Career	Number in Group	Average Distance	Ideal Type		
Organists:					
Court organists	4	.226	40CRTORG		
Church organists	11	.110	40CHUORG		
Court system:					
Kapellmeisters:					
Kapellmeister	9	.181	45CRTKPM		
Kapellmeister/other	2	.164	12CRTKPM	35OTHER	
Kapellmeister/other/kapellmeister	2	.312	8CRTKPM	22OTHER	18CRTKPM
Vocalists:					
Vocalist/kapellmeister	4	.424	10CRTVOC	33CRTKPM	
Instrumentalists/konzertmeisters:					
Instrumentalist	17	.037	48CRTINS		
Konzertmeister	4	.014	50CRTKZM		
Short-term instrumentalist/konzertmeister	2	.418	8CRTINS	27CRTKZM	

Long-term instrumentalist/konzertmeister	4	.128	35CRTINS	12CRTKZM	
(Other court?)/administrator	5	.486	45CRTADM		
Instrumentalist/konzertmeister/kapellmeister	5	.292	10CRTINS	7CRTKZM	29CRTKPM
Other court positions:					
Court composer/other	2	.382	16CRTCOM	16OTHER	
Church system:					
Church kapellmeister	2	.020	50CHUKPM		
Church cantor	3	.425	25CHUCAN		
Church administrator	2	.020	50CHUADM		
Other jobs:					
Long-term other	4	.112	45OTHER		
Other/kapellmeister	4	.324	35OTHER	10CRTKPM	
Unclassified	6	.810	Other		
Average within groups		.137			
Average between groups		.810			
Jackknife ratio estimate		1.901			
t-value		13.17			
df		91			
Stress statistic		.157			

TABLE 4

CAREER PATTERNS OF GERMAN MUSICIANS: LOGARITHMIC MEASURE

Type of Career	Number in Group	Average Distance	Ideal Type	
Organists:				
Long-term court organists	6	.528	9CRTORG	
Long-term church organists	8	.074	11CHUORG	
Other/church organists	4	.358	5OTHER	9CHUORG
Court system:				
Kapellmeisters:				
Medium-term kapellmeisters	7	.249	9CRTKPM	
Kapellmeister/other	2	.394	8CRTKPM	5OTHER
Vocalists:				
Vocalist/kapellmeister	2	.363	5CRTVOC	10CRTKPM
Vocalist/konzertmeister	4	.520	6CRTVOC	11CRTKZM
Instrumentalists/konzertmeisters:				
Short-term instrumentalist	3	.121	5CRTINS	
Medium-term instrumentalist	15	.195	10CRTINS	
Short-term instrumentalist/konzertmeister	6	.288	4CRTINS	10CRTKZM

Instrumentalist/konzertmeister/administrator	3	.386	3CRTINS	7CRTKZM	6CRTADM
Any court/other/konzertmeister	2	.556	5CRTOTH	5OTHER	8CRTKZM
Instrumentalist/kapellmeister	3	.356	6CRTINS	9CRTKPM	
Any court/kapellmeister	2	.519	8CRTOTH	8CRTKPM	
Other court positions:					
Court composer	3	.458	5CRTCOM		
Church system:					
Church kapellmeister	2	.699	7CHUKPM		
Church cantor	2	.208	8CHUCAN		
Opera system:					
Opera kapellmeister	2	.578	8OPEKAP		
Town system:					
Town instrumentalist/town administrator	3	.686	5TWNINS	6TWNBAN	
Other jobs:					
Unclassified	14	.849	Other		
Average within groups		.387			
Average between groups		.825			
Jackknife ratio estimate		1.434			
t-value		4.57			
df		92			
Stress statistic		.156			

one lose, in this study, the full sequence information that is neglected in most methods based on stochastic modeling of individual transitions. In fact, one retains a focus on the whole structure of social narratives that is essential to theoretical writers as disparate as Weber and the symbolic interactionists. Optimal matching also permits the analysis of nonstochastic sequence data—ritual sequences, labor process sequences—which standard methods cannot consider.

The strengths of optimal matching can be realized, however, only with clear recognition of its assumptions and careful specification of its parameters. Thus, like all methods analyzing careers individually, optimal matching must assume that the sequences studied are independent. While the results here strongly point to a set of typical careers, one cannot accept that verdict without an analysis of intercareer contingency by other methods, such as vacancy-chain analysis. Were vacancy-chain analysis to show strong determination of musicians' careers by vacancies and opportunities, then the typicality of careers would require reconceptualization as a typicality of vacancy patterns.

Optimal matching also requires explicit parameterization of substitution, since substitution costs may affect distances. Finding adequate theoretical grounds and empirical measures for the resemblance of sequence elements is a central and difficult task. However, Abbott and Forrest's (1988) experimental work indicates that the difficulties are not massive. There is substantial robustness with respect to substitution costs and even with respect to event coding itself. To be sure, their judges agreed on the overall nature of the task and on certain general patterns of resemblance among the events, disagreeing mostly about finer distinctions. But even those conclusions based on secondary analyses—the clusterings and scalings based on the distance matrices produced by optimal matching—remained the same across the judges. Thus, while substitution must be carefully handled, it is not a supersensitive task whose errors will be compounded by later stages in the analysis.[21]

[21] In standard methods, there are several large literatures on the issues of conceptualization, measurement, and analysis that consider the many effects of early decisions on later analysis. Since nonstochastic study of sequences is relatively unusual, there is no such literature. (However, Abbott [1984] begins an analysis on the conceptualization and measurement of sequence data.) The worry that different measurement and analytic assumptions lead to different results arises with any method, a fact evident in the commentary sections of the *American Journal of Sociology* and the *American Sociological Review*. The exchanges between Hodson and Kaufman (1981) and Horan, Tolbert, and Beck (1981) on the dual labor market, between Hartman and Hsiao (1988), Weede (1985), and Muller (1985, 1988) on inequality and violence, and between Slomczynski (1984) and Krause and Sobel (1984) on types of mobility are merely a few of many possible examples.

It is important to recall, too, that these dangers are dangers common to virtually all sociological methods. Every method makes assumptions about the measurement of variables, about the meaning of time, about the independence of cases; optimal matching does not make unusually strong ones. Optimal matching also resembles some more familiar techniques in its reliance on fairly indirect statistical tests (here the jackknife test of the groups), a reliance it shares with most Markovian models for studying mobility. One tests vacancy-chain models, for example, by fitting distributions of chain length.

A number of readers have questioned the technique on the grounds that it involves two phases, a data-reduction phase (the optimal matching proper) and a further analysis phase (the clustering or scaling). But here, too, optimal matching techniques resemble more familiar methods. This two-phase quality is characteristic of fixed measurement-model LISREL as of any application of multidimensional scaling. In the one case, factor analysis is followed by structural models among the factors; in the other, a distance model is followed by the iterative optimization of an objective function based on that distance. In both cases, many analytic decisions are made on theoretical and practical grounds in the data-reduction phase that have varying and often unexplored effects on actual analysis downstream. That these decisions now have relatively conventional answers (e.g., that Jaccard distance is preferred in some scaling applications and matching distance in others) reflects the maturity of the techniques involved. Time and use will undoubtedly produce similar conventions for optimal matching. It is thus premature to criticize the technique on the ground that arbitrary decisions cumulate throughout the process to undermine results.

Like any method, however, optimal matching has certain weaknesses. As a method of direct analysis, it requires dyadic data, a fact that severely constrains data set size. And optimal matching is not a substitute for, but a complement to, stochastic analysis. Careers and similar processes do unfold in time, and predicting their development at any given point in time requires probabilistic modeling embedded in a directional time, as in many of the formal career models mentioned in Section I. Optimal matching may be useful in reducing a mass of sequences to a small number of patterns, but one must still employ stochastic methods (e.g., event-history methods, Markov models generally) to find the forces producing those patterns. (Where stochastic modeling is not appropriate and only sequence resemblance is at issue, however, optimal matching is our only recourse.) Finally, the methods are not yet widely available in software designed for social science applications. Most biological packages have them but, being designed for DNA analysis, lack the kinds of input/

output facilities sociologists want. However, we hope in this article to kindle an interest that will lead to development of such techniques in standard packages.

In summary, we have here presented a simple and widely used technique for measuring sequence resemblance, showing how it can effectively analyze a typical sociological data set involving sequence questions. The test of such a method lies in its use. We feel that it has produced a far more effective categorization of careers in our data than could have been produced by any other method currently available. We encourage others to test it and hope it can find general use in the discipline.

APPENDIX

The problem of minimizing intersequence distance subject to varying costs of substitution is solved by an iterative minimization procedure. Figure A1 illustrates the transformation of sequence A (1 1 2 3 4 5 6) into sequence D (4 5 6 7 7 7 7 7). The figure is a matrix with one more row than one of the sequences (A) and one more column than the other (D). One sequence (by convention the top one in an alignment, here D) spans the later columns of the matrix, leaving the first column without a label. The other sequence (the bottom in the alignment, here A) spans the later rows, leaving the first row without a label. Note that "interior" cells have four elements, while the top and left border cells have two, except for the upper left-hand corner cell, which has only one. The reasons for this will be clear shortly.

To create the minimum cost alignment we move through this matrix downward and rightward. (This procedure is equivalent to building an alignment from left to right.) Moving right one column means "adding an insertion" to the alignment. (The choice of the word "insertion" here is conventional.) This means putting the column label in the top member of the alignment and ϕ (for null) in the bottom member. Moving down one row means "adding a deletion," which conversely means putting the row label in the bottom member of the alignment and ϕ in the top member. Moving down one row *and* to the right one column (i.e., diagonally downward and rightward) means "adding a substitution," putting the column label in the top member of the alignment and the row label in the bottom member. The costs of these motions are recorded in the matrix; insertion cost is in the lower left of each cell, deletion cost is in the upper right, substitution cost in the upper left. For the present example, we have assumed that insertion and deletion cost one unit each and that substitution of job-level i for job-level j costs $|i - j|/10$, where i and j are drawn from $\{1, 2, \ldots, 10\}$.

The lower right element in each cell is the minimum cost of arriving

		4	5	6	7	7	7	7	7
	0.0	1.0 1.0	1.0 2.0	1.0 3.0	1.0 4.0	1.0 5.0	1.0 6.0	1.0 7.0	1 0 8.0
1	1.0 1.0	0.3 1.0 1.0 0.3	0.4 1.0 1.0 1.3	0.5 1.0 1.0 2.3	0.6 1.0 1.0 3.3	0.6 1.0 1.0 4.3	0.6 1.0 1.0 5.3	0.6 1 0 1.0 6 3	0.6 1 0 1.0 7 3
1	1.0 2.0	0.3 1.0 1.0 1.3	0.4 1.0 1.0 0.7	0.5 1.0 1.0 1.7	0.6 1.0 1.0 2.7	0.6 1.0 1.0 3.7	0.6 1.0 1.0 4.7	0.6 1.0 1.0 5.7	0.6 1.0 1.0 6.7
2	1.0 3.0	0.2 1.0 1.0 2.2	0.3 1.0 1.0 1.6	0.4 1.0 1.0 1.1	0.5 1.0 1.0 2.1	0.5 1.0 1.0 3.1	0 5 1.0 1.0 4.1	0.5 1 0 1.0 5.1	0.5 1 0 1.0 6 1
3	1.0 4.0	0.1 1.0 1.0 3.1	0.2 1.0 1.0 2.4	0.3 1.0 1.0 1.9	0.4 1.0 1.0 1.5	0.4 1.0 1.0 2.5	0.4 1.0 1.0 3.5	0.4 1.0 1.0 4.5	0.4 1.0 1 0 5 5
4	1.0 5.0	0.0 1.0 1.0 4.0	0.1 1.0 1.0 3.2	0.2 1.0 1.0 2.6	0.3 1.0 1.0 2.2	0.3 1.0 1.0 1.8	0.3 1.0 1.0 2.8	0.3 1.0 1.0 3.8	0.3 1.0 1.0 4 8
5	1.0 6.0	0.1 1.0 1.0 5.0	0.0 1.0 1.0 4.0	0.1 1.0 1.0 3.3	0.2 1.0 1.0 2.8	0.2 1.0 1.0 2.4	0.2 1.0 1.0 2.0	0.2 1.0 1.0 3.0	0.2 1.0 1.0 4 0
6	1.0 7.0	0.2 1.0 1.0 6.0	0.1 1.0 1.0 5.0	0.0 1.0 1.0 4.0	0.1 1.0 1.0 3.4	0.1 1.0 1.0 2.9	0.1 1.0 1.0 2.5	0.1 1.0 1.0 2 1	0.1 1.0 1 0 3.1

FIG. A1.—Sequence 1: 4 5 6 7 7 7 7 7; sequence 2: 1 1 2 3 4 5 6

there. Hence, the cost at the start, in the upper left corner of the matrix, is zero (and, since we start there, the zero is the only cell entry). Since we move rightward along the first row only by continuously "inserting" (in the sense just defined), there are only insertion costs and total costs in these border cells. Each additional column costs one, and we end up, at the upper right-hand corner, with a cost of 8; we now have sequence D on top in the alignment and a row of ϕ's on the bottom. Similarly, since we move down the first *column* only by continuous "deletion," each additional row costs us one, and we end up at the lower left corner with a cost of 7; here we have sequence A on the bottom in the alignment, and 7 ϕ's on top.

Since we wish to arrive at a full alignment of both sequences, we must arrive at the lower right corner. Note that any interior cell can be entered three ways: from the left (by insertion), from above (by deletion), and diagonally from left above (by substitution). Assuming we have already calculated the minimum costs of arriving at these prior cells, we simply add these prior costs to the cost of entering the present cell from each of those cells and choose the minimum of these sums for the lower right of the present cell. Thus we must first calculate the uppermost left interior cell (what would be cell 2,2 with normal matrix labels), because only for it do we possess the three prior figures. We can enter from left at cost of one or from above at cost of one. In each case the prior cost was one and the cost of entry is one, for a total of two. We can also enter diagonally, with prior cost zero and entry cost 0.3. Clearly 0.3 is the minimum total cost; it hence becomes the lower right entry of the cell. Each of the three paths embodies a different alignment:

	ϕ 4		4 ϕ		4
From the left:		From above:		Diagonally:	
	1 ϕ		ϕ 1		1

Given our current model for substitution costs, the last is clearly the minimum cost move. We proceed similarly through the matrix downward and rightward until we reach the answer 3.1. Divided by the length of the longer sequence, this produces 0.388. Note that this result can be reached in a total of five different ways, depending on where one places the required insertion. The arrows on figure 1 mark the possible pathways.

The monotonic nature of these sequences makes the solution appear deceptively simple. Figure A2 presents a quite unlikely pair of career paths to show how the method works with less obviously ordered data. Here the issue is whether to shift the upper sequence to the right to pick up the 1–9–8 match at the beginning or to the left to pick up the 8–9–1 match at the end. The comparative costs are 3.6 and 3.5. The extreme character of these sequences (they involve only one pair of elements with small interelement distance) indicates something important about the method; if substitution costs are low relative to insertion and deletion costs, the algorithm will always choose substitution. Substitution costs must therefore be carefully scaled to fit the particular application.

		1	9	8	1	8	9	1
	0.0	1.0 1.0	1.0 2.0	1.0 3.0	1.0 4.0	1.0 5.0	1.0 6.0	1.0 7.0
8	1.0 1.0	0.7 1.0 1.0 0.7	0.1 1.0 1.0 1.1	0.0 1.0 1.0 2.0	0.7 1.0 1.0 3.0	0.0 1.0 1.0 4.0	0.1 1.0 1.0 5.0	0.7 1.0 1.0 6.0
1	1.0 2.0	0.0 1.0 1.0 1.0	0.8 1.0 1.0 1.5	0.7 1.0 1.0 1.8	0.0 1.0 1.0 2.0	0.7 1.0 1.0 3.0	0.8 1.0 1.0 4.0	0.0 1.0 1.0 5.0
9	1.0 3.0	0.8 1.0 1.0 2.0	0.0 1.0 1.0 1.0	0.1 1.0 1.0 1.6	0.8 1.0 1.0 2.6	0.1 1.0 1.0 2.1	0.0 1.0 1.0 3 0	0.8 1.0 1.0 4.0
8	1.0 4.0	0.7 1.0 1.0 3.0	0.1 1.0 1.0 2.0	0.0 1.0 1.0 1.0	0.7 1.0 1.0 2.0	0.0 1.0 1.0 2.6	0.1 1.0 1.0 2.2	0 7 1.0 1 0 3 2
9	1.0 5.0	0.8 1.0 1.0 4.0	0.0 1.0 1.0 3.0	0.1 1.0 1.0 2.0	0.8 1.0 1.0 1.8	0.1 1.0 1.0 2.1	0.0 1.0 1.0 2.6	0.8 1.0 1.0 3.0
1	1.0 6.0	0.0 1.0 1.0 5.0	0.8 1.0 1.0 4.0	0.7 1.0 1.0 3.0	0.0 1.0 1.0 2.0	0.7 1.0 1.0 2.5	0.8 1.0 1.0 2 9	0.0 1 0 1.0 2.6
9	1.0 7.0	0.8 1.0 1.0 6.0	0.0 1.0 1.0 5.0	0.1 1.0 1.0 4.0	0.8 1.0 1.0 3.0	0.1 1.0 1.0 2.1	0.0 1.0 1.0 2.5	0.8 1.0 1.0 3 5

FIG. A2.—Sequence 1: 1 9 8 1 8 9 1; sequence 2: 8 1 9 8 9 1 9

Optimal matching as described here is easily programmed. Some elaborations, in particular those using transposition, reversal, expansion, and contraction as elementary operations like insertions, deletions, and substitutions, require more substantial effort. Algorithms (but not the actual computer code) are provided for most of these elaborations in Sankoff and Kruskal (1983). Most commercially available DNA-sequencing programs have optimal matching routines (see, e.g., Pustell and Kafatos 1986); the standard biological algorithms are those of Wilbur and Lipman (1982, 1984; Lipman and Pearson 1984). Beyond the many examples in Sankoff and Kruskal (1983), recent developments are covered in Miura (1986) and formal properties are analyzed in Day (1984). Our calculations were all done using the Beldings Program Series, graciously provided by David Bradley of the California State University at Long Beach.

Optimal matching in the simple form discussed here does not require massive computer resources. Execution times are roughly proportional to (*a*) the square of the length of the longest sequence and (*b*) the square of the number of sequences analyzed. In sociological applications the first of these is unlikely to be large. Biological applications often work with sequences of thousands of proteins and must use special search techniques (hashing methods) to speed the process. More problematic for sociologists is the second factor, the data set size. Here the constraints are much the same as with other methods that use relational data (data defined for every pair of units of analysis), of which the most familiar is network analysis. Data sets of under 100 items are quickly analyzed, data sets of several hundred items are practical but analyzed somewhat more slowly, and data sets of several thousand items are impractical. We discuss this limitation in text and suggest a method for avoiding it (an extensive discussion of computer time and space issues may be found in Sankoff and Kruskal [1983], pt. 4).

One reader of this article has raised the issue of censoring. Optimal matching methods as currently used assume all analyzed sequences to be equally complete. If the degree of censoring is known case by case, the methods can easily be adapted to function effectively. Where degree of censoring is unknown, they could in principle be programmed to estimate that degree of censoring on certain assumptions about underlying sequence equivalence.

REFERENCES

Abbott, A. 1983. "Sequences of Social Events." *Historical Methods* 16:129–47.

———. 1984. "Event Sequence and Event Duration." *Historical Methods* 17:192–204.

Abbott, A., and J. Forrest. 1986. "Optimal Matching for Historical Sequences." *Journal of Interdisciplinary History* 16:471–94.

———. 1988. "The Optimal Matching Method for Studying Anthropological Data: An Introduction and Reliability Analysis." Unpublished manuscript. Rutgers University, Department of Sociology.

Abell, P. 1984. "Comparative Narratives." *Journal for the Theory of Social Behavior* 14:309–31.

———. 1985. "Analyzing Qualitative Sequences." Pp. 99–115 in *Sequence Analysis,* edited by P. Abell and M. Proctor. Brookfield, Vt.: Gower.

———. 1987. *The Syntax of Social Life.* Oxford: Oxford University Press.

Becker, H. S. 1952. "The Career of the Chicago Public School Teacher." *American Journal of Sociology* 57:470–77.

Bertaux, D. 1982. "The Life Course Approach as a Challenge to the Social Sciences." Pp. 127–50 in *Aging and Life Course Transitions,* edited by T. K. Hareven and K. J. Adams. New York: Guilford.

Blume, F. 1949–79. *Die Musik in Geschichte und Gegenwart,* 17 vols. Kassel: Bärenreiter.

Blumstein, A., and J. Cohen. 1987. "Characterizing Criminal Careers." *Science* 237: 985–91.

Boudon, R. 1973. *Mathematical Structures of Social Mobility.* San Francisco: Jossey-Bass.

Campbell, R. L., and D. M. Richie. 1983. "Problems in the Theory of Developmental Sequences." *Human Development* 26:156–72.

Carroll, G. R., and K. V. Mayer. 1986. "Job Shift Patterns in the Federal Republic of Germany." *American Sociological Review* 51:323–41.

Chalmers, B. E. 1981. "A Selective Review of Stress." *Current Psychological Reviews* 1:325–43.

D'Amico, R. 1985. "The Effect of Career Origins on Subsequent Socio-economic Achievement." *Work and Occupations* 12:329–50.

D'Amico, R., and R. T. Brown. 1982. "Patterns of Labor Mobility in a Dual Economy." *Social Science Research* 11:153–75.

David, P. A. 1985. "Clio and the Economics of QWERTY." *American Economic Review* 75:332–37.

Day, W. H. E. 1984. "Properties of Levenshtein Metrics on Sequences." *Bulletin of Mathematical Biology* 46:327–32.

Dewatripont, M. 1987. "The Role of Indifference in Sequential Models of Spatial Competition." *Economics Letters* 23:323–28.

Dietterich, T. G., and R. S. Michalski. 1985. "Discovering Patterns in Sequences of Events." *Artificial Intelligence* 25:187–232.

Dix, G. 1945. *The Shape of the Liturgy.* London: Dacre.

Evans, M. D., and E. O. Laumann. 1983. "Professional Commitment." Pp. 3–40 in *Social Stratification and Mobility,* vol. 2. Greenwich, Conn.: JAI.

Featherman, D. L. 1971. "A Research Note: A Social Structural Model for the Socioeconomic Career." *American Journal of Sociology* 77:293–304.

Form, W. H., and D. G. Miller. 1949. "Occupational Career Patterns as a Sociological Instrument." *American Journal of Sociology* 54:317–29.

Gaertner, K. N. 1980. "The Structure of Organizational Careers." *Sociology of Education* 53:7–20.

Glick, P. C. 1976. "Updating the Life Cycle of the Family." *Journal of Marriage and the Family* 39:5–13.

Grandjean, B. D. 1981. "History and Career in a Bureaucratic Labor Market." *American Journal of Sociology* 86:1057–92.

Hall, O. 1948. "The Stages of a Medical Career." *American Journal of Sociology* 53:327–36.

Hannan, M. T., and N. B. Tuma. 1984. *Social Dynamics.* New York: Academic Press.

Hartigan, J. A. 1975. *Clustering Algorithms*. New York: Wiley.
Hartman, J., and W. Hsiao. 1988. "Inequality and Violence: Issues of Theory and Measurement in Muller." *American Sociological Review* 53:794–99.
Heckman, J. J. 1978. "Life Cycle and Household Decision Making." *American Economic Review* 68:200–207.
Heckman, J. J., and T. E. MaCurdy. 1980. "A Life-Cycle Model of Female Labour Supply." *Review of Economic Studies* 47:47–74.
Hodson, F. R., D. G. Kendall, and P. Tautu, eds. 1971. *Mathematics in the Archeological and Historical Sciences*. Edinburgh: University of Edinburgh.
Hodson, R. D., and R. L. Kaufman. 1981. "Circularity in the Dual Economy." *American Journal of Sociology* 86:881–87.
Hogan, D. P. 1978. "The Variable Order of Events in the Life Course." *American Sociological Review* 43:573–86.
Hogwood, C. 1977. *Music at Court*. London: Folio Society.
Hopp, W. J. 1987. "A Sequential Model of R and D Investment over an Unbounded Time Horizon." *Management Science* 33:500–508.
Horan, P. M., C. M. Tolbert, and E. M. Beck. 1981. "The Circle Has No Close." *American Journal of Sociology* 86:887–94.
Hrycak, A. 1988. "Mobility and Music in Eighteenth Century Germany." Henry Rutgers Honors Thesis. Rutgers University.
Hubert, L. J. 1976. "Seriation Using Asymmetric Proximity Measures." *British Journal of Mathematical and Statistical Psychology* 29:32–53.
———. 1979. "Comparison of Sequences." *Psychological Bulletin* 86:1098–1106.
Kalleberg, A. L., and P. M. Hudis. 1979. "Wage Change in the Late Career." *Social Science Research* 8:16–40.
Kelley, J. 1973. "Causal Chain Models for the Socioeconomic Career." *American Sociological Review* 38:481–93.
Konda, S. L., and S. Stewman. 1980. "An Opportunity Labor Demand Model and Markovian Labor Supply Models." *American Sociological Review* 45:276–301.
Kruskal, J. B., and M. Liberman. 1983. "The Symmetric Time-Warping Problem." Pp. 125–61 in *Time Warps, String Edits, and Macromolecules,* edited by D. Sankoff and J. B. Kruskal. Reading, Mass.: Addison-Wesley.
Levinson, D. J. 1978. *The Seasons of a Man's Life*. New York: Knopf.
Lipman, D. J., and W. R. Pearson. 1984. "Rapid and Sensitive Protein Similarity Searches." *Science* 227:1435–41.
Lortie, D. C. 1959. "Laymen to Lawmen." *Harvard Educational Review* 29:352–69.
MaCurdy, T. E. 1981. "An Empirical Model of Labor Supply in a Life Cycle Setting." *Journal of Political Economy* 89:1059–85.
March, J. C., and J. G. March. 1977. "Almost Random Careers." *Administrative Science Quarterly* 22:377–409.
McClure, E. 1983. "Sociocultural Variables in Children's Sequencing of Stories." *Discourse Processes* 6 (2): 131–43.
Miura, R. M. 1986. *Some Mathematical Questions in Biology: DNA Sequence Analysis*. Vol. 17 of *Lectures on Mathematics in the Life Sciences*. Providence: American Mathematical Society.
Muller, E. N. 1985. "Income Inequality and Political Violence: The Effect of Influential Cases." *American Sociological Review* 51:441–45.
———. 1988. "Inequality, Repression and Violence: Issues of Theory and Research Design." *American Sociological Review* 53:799–806.
Oppenheimer, L., and M. Groot. 1985. "Construction and Reconstruction of Social Event Sequences." *Psychological Reports* 56:511–18.
Oppenheimer, L., and H. Van der Lee. 1983. "Social Perspective-Taking and Event Sequences." *Psychological Reports* 53:683–90.
Poole, M. S. 1983. "Decision Development in Small Groups. III. A Multiple Sequence

Model of Group Decision Development." *Communication Monographs* 50:321–41.

Pustell, J., and F. C. Kafatos. 1986. "A Convenient and Adaptable Microcomputer Environment for DNA and Protein Sequence Manipulation and Analysis." *Nucleic Acids Research* 14:477–88.

Raynor, H. 1972. *A Social History of Music from the Middle Ages to Beethoven.* London: Barrie & Jenkins.

Reissman, L. 1956. "Life Careers, Power, and the Professions: The Retired Army General." *American Sociological Review* 21:215–21.

Rosenbaum, J. E. 1979*a*. "Tournament Mobility." *Administrative Science Quarterly* 24:220–41.

———. 1979*b*. "Organizational Career Mobility." *American Journal of Sociology* 85:21–48.

———. 1984. *Career Mobility in a Corporate Hierarchy*. Orlando, Fla.: Academic.

Rosenfeld, R. A. 1979. "Women's Occupational Careers." *Sociology of Work and Occupations* 6:283–311.

———. 1980. "Race and Sex Differences in Career Dynamics." *American Sociological Review* 45:583–609.

Sadie, S., ed. 1980. *The New Grove Dictionary of Music and Musicians,* 20 vols. London: Macmillan.

Salmen, W., ed. 1983. *The Social Status of the Professional Musician from the Middle Ages to the 19th Century*. New York: Pendragon.

Sankoff, D., and J. B. Kruskal, eds. 1983. *Time Warps, String Edits, and Macromolecules*. Reading, Mass.: Addison-Wesley.

Singer, B., and S. Spilerman. 1976. "The Representation of Social Processes by Markov Models." *American Journal of Sociology* 82:1–54.

Skocpol, T. 1979. *States and Social Revolutions*. Cambridge: Cambridge University Press.

Skvoretz, J. 1984*a*. "Career Mobility as a Poisson Process." *Social Science Research* 13:198–220.

———. 1984*b*. "The Logic of Opportunity and Mobility." *Social Forces* 63:72–97.

Slomczynski, K. M., and T. Krause. 1984. "Should the Framework of Structural versus Circulation Mobility Be Abandoned?" *American Sociological Review* 49:850–52.

Smigel, E. O. 1964. *The Wall Street Lawyer*. New York: Free Press.

Sobel, M. E. 1984. " 'Frameworks' and Frameworks." *American Sociological Review* 49:852–53.

Sørensen, A. B. 1974. "A Model for Occupational Careers." *American Journal of Sociology* 80:44–57.

Spanier, G. B., W. Sauer, and R. Larzelere. 1979. "An Empirical Evaluation of the Family Life Cycle." *Journal of Marriage and the Family* 41:27–38.

Spilerman, S. 1977. "Careers, Labor Market Structure, and Socioeconomic Achievement." *American Journal of Sociology* 83:551–93.

Stewman, S. 1975. "Two Markov Models of Open System Occupational Mobility." *American Sociological Review* 40:298–321.

———. 1976. "Markov Models of Occupational Mobility." *Journal of Mathematical Sociology* 4:201–45, 247–78.

———. 1986. "Demographic Models of Internal Labor Markets." *Administrative Science Quarterly* 31:212–47.

Stewman, S., and S. L. Konda. 1983. "Careers and Organizational Labor Markets." *American Journal of Sociology* 88:637–85.

Stinchcombe, A. L. 1978. *Theoretical Methods in Social History*. New York: Academic Press.

Tolbert, C. M. 1982. "Industrial Segmentation and Men's Career Mobility." *American Sociological Review* 47:457–77.

Treisman, M., and T. C. Williams. 1984. "A Theory of Criterion Setting with an Application to Sequential Dependencies." *Psychological Review* 91:68–111.

Vroom, V. H., and K. R. MacCrimmon. 1968. "Toward a Stochastic Model of Managerial Careers." *Administrative Science Quarterly* 13:1326–46.

Vuchinich, S. 1984. "Sequencing and Social-Structure in Family Conflict." *Social Psychology Quarterly* 47:217–34.

Wanner, R. A., and L. S. Lewis. 1983. "Economic Segmentation and the Course of the Occupational Career." *Work and Occupations* 10:307–24.

Weber, M. 1949. *The Methodology of the Social Sciences,* translated by E. A. Shils and H. A. Finch. Glencoe, Ill.: Free Press.

Weede, E. 1985. "Income Inequality and Political Violence Reconsidered." *American Sociological Review* 51:438–41.

White, H. C. 1970. *Chains of Opportunity*. Cambridge, Mass.: Harvard University Press.

Wilbur, W. J., and D. J. Lipman. 1982. "Rapid Similarity Searches of Nucleic Acid and Protein Data Banks." *Proceedings of the National Academy of Sciences* 80:726–30.

———. 1984. "The Context Dependent Comparison of Biological Sequences." *SIAM Journal of Applied Mathematics* 44:557–67.

Wilensky, H. C. 1961. "Orderly Careers and Social Participation." *American Sociological Review* 26:521–39.

———. 1964. "The Professionalization of Everyone?" *American Journal of Sociology* 70:137–58.

Wyer, R. S., E. J. Shoben, R. W. Fuhrman, and G. V. Bodenhausen. 1985. "Event Memory: The Temporal Organization of Social-Action Sequences." *Journal of Personality and Social Psychology* 49:857–77.