

Typologizing Temporality: Time-Aggregated and Time-Patterned Approaches to Conceptualizing Homelessness

William McAllister
Columbia University

Li Kuang
Columbia University

Mary Clare Lennon
City University of New York

This article employs a relatively new method to construct time-based typologies of homelessness, arguing that time-aggregated typologies in previous research lose useful information by summing, averaging, or otherwise summarizing events that occur over time. This study instead proposes a time-patterned approach that measures the timing, duration, and sequence of events as they unfurl over time. It first compares the two approaches by examining support for a theorized three-category typology analyzed by Randall Kuhn and Dennis Culhane. Both approaches identify the three groups initially found by Kuhn and Culhane, but the time-patterned approach performs marginally better. Both analyses leave too much heterogeneity in the groups, and the initial theory for the three categories is not robust. These deficiencies suggest the utility of further analysis. Using a time-patterned analysis, this study then identifies 10 temporally based homeless groups that strongly differ from the three groups found by Kuhn and Culhane. It then organizes these 10 groups into four sets of groups and speculates about how structural factors and individual traits can combine to generate these categories.

Researchers improve theories and policy making by forming better concepts. Forming typologies helps researchers create such concepts. This article uses a relatively new method to construct time-based homeless typologies that expand the ability to theorize and to make policy. The

key argument is that commonly used time-aggregated typologies lose potentially useful information by summing, averaging, or otherwise summarizing events that occur over time. This study instead suggests a time-patterned approach that captures events as they unfurl over time by measuring the timing, duration, and sequence of homelessness.

This article begins by describing the prevailing three-category typology in homeless research and policy making. It explains some problems that stem from how this typology employs temporal information, and it describes a time-patterned approach that allows researchers to utilize more of that information. The study then carries out a time-patterned analysis of the typology and compares its results with those from a time-aggregated analysis. The study finds that the time-patterned approach performs marginally better. Another finding is more important, however. The extent of within-group variability in the results of both approaches and the weakness of theorized relations suggest that the prevailing three-category typology can be improved. Further analysis of time-patterned results identifies 10 groups that technically and substantively contrast strongly with the three groups found by Randall Kuhn and Dennis Culhane (1998). The study then organizes these 10 groups into sets and uses these sets to suggest how focusing on temporal change allows researchers to generate theories about the interaction of structural factors and individual traits for generating homelessness.

Homeless Typologizing

Typologies combine different values of relevant phenomena to identify concepts such as the chemical elements or Weberian types of organizations (Stinchcombe 1968; Bailey 1994; Doty and Glick 1994). Randall Kuhn and Dennis Culhane (1998) report the major theory on and develop evidence for a typology that reflects the temporal experience of homelessness among adults unaccompanied by mates or children.¹ Based on extant homelessness research, they characterize three types of homelessness—the transitionally homeless, the episodically homeless, and the chronically homeless.² In the context of homeless shelter use, Kuhn and Culhane describe the transitionally homeless as people who “are forced to spend a short time in a homeless shelter before making a transition into a more stable housing arrangement, and in most cases they do not return to homelessness” (1998, 211). Episodically homeless people are described as those who “frequently shuttle in and out of homelessness. . . . [They] often find their way back to the shelters” (1998, 211). Kuhn and Culhane describe the chronically homeless as people “likely to be entrenched in the shelter system, . . . for whom shelters are more like long-term housing than an emergency arrangement” (1998, 211). The transitional category is expected to be much

larger than the other two, which are expected to be about the same size.

To allow them to test for further support for this typology, Kuhn and Culhane (1998) extract from then-extant homelessness research theory about the relationship between each kind of homelessness and individuals' features. Transitionally homeless people are thought to be younger than those chronically homeless and to have fewer mental health, substance abuse, or other medical problems compared with those episodically or chronically homeless. Episodically homeless people are thought to be younger than those who are chronically homeless and to have a greater likelihood of more medical, mental health, and substance abuse problems than the transitionally homeless. They are also thought to have a greater likelihood of mental health and substance abuse problems than those chronically homeless. Last, chronically homeless people are expected to be relatively older and to have more medical, mental health, and substance abuse problems than those who are transitionally homeless and more medical problems than those who are episodically homeless.

To empirically analyze this theory, Kuhn and Culhane (1998) use two temporal dimensions to construct the typology: the frequency and the duration of homelessness. They utilize administrative shelter data to measure these dimensions for first-time sheltered people by counting the number of shelter episodes (frequency) and the total number of days sheltered (duration) over 3 years. They posit that transitionally homeless people enter shelters once or twice for a very short time, that episodically homeless people enter shelters many times but for short periods, and that chronically homeless people enter shelters once or very few times but spend most of their time homeless.

Kuhn and Culhane find the expected number of groups, their expected sizes, and their expected natures. They also show that traits hypothesized to be associated with each group are generally associated as posited and that the relations are in the right direction. (See Kertesz et al. 2005 for further support regarding hypothesized traits.)

Time-Aggregated and Time-Patterned Approaches

In Kuhn and Culhane's (1998) approach, time is aggregated by treating each occurrence of an event as a cross-sectional variable. They sum each person's total number of homeless days or episodes over a specified period. Research in this approach may also calculate rates, averages, or proportions, or may otherwise aggregate events over time, and it may do so across or within individuals. Such time-aggregated measures, however, cannot represent patterns of homelessness or housing situations that occur over time in people's lives. That is, a typology might be more faithful to how people experience their lives and more analytically useful

Persons	Time Periods																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
B	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

FIG. 1.—Hypothetical housing histories over 24 time periods. Note: Black cells indicate that the person was homeless for that entire period; blank cells indicate that the person was housed for that entire period.

if it is based on patterns that capture when, for how long, and in what sequence homelessness occurs as people's lives unfold. This point is illustrated in figure 1, which shows housing histories for two hypothetical persons over 24 time periods. A blank cell indicates that the person was housed for the entire time period, and a black cell indicates that he or she was homeless for the entire period.

If one takes a time-aggregated approach, the homeless histories of these two people appear to be identical. They are homeless for the same number of days and have the same number of homeless episodes. However, if one takes a time-patterned approach that analyzes the timing, duration, and sequence of homeless and nonhomeless episodes, the two individuals have very different histories. This analysis suggests that individual A perhaps fitfully frees herself or himself from homelessness while individual B tumbles further into it. As a result, and in contrast to the time-aggregated analysis, researchers might theorize differently about these peoples' lives, as they would seek to explain different histories or hypothesize that these different histories have different effects.

In contrast to typologies produced in a time-aggregated way, a time-patterned approach allows researchers to construct typologies that identify more refined concepts regarding the temporal character of homelessness. Relative to time-aggregation, time-patterning expands the number of dimensions observed (timing and sequence of homeless and nonhomeless episodes) and effectively translates frequency and duration into more subtle measures by specifying when each episode occurs and how long each lasts. The more analytically useful dimensions are employed and the more those dimensions are subtly measured, the richer typology-based concepts can be. Incorporating how events occur over time therefore seems an appropriate way to refine prevailing theorizing on the temporal character of homelessness. Time-patterning allows researchers to see, for example, the ambiguity in the meaning of a "relatively short" episode and the length of time between episodes (Kuhn and Culhane's description of the transitional category; 1998, 211). It also allows constructing a typology that does not require relying on such characterizations of time but rather incorporates whatever the time interval is. Time-patterning allows researchers to see and empirically address similar ambiguities better than does time aggregation.

The remainder of this article supports these arguments by using both approaches to analyze homeless shelter data. It first compares results from the two analyses for the theorized three-category typology and finds that the time-patterned analysis produces marginally better results than those from the time-aggregated analysis. However, within-group heterogeneity remains great in both analyses, and the time-patterned results are not technically optimal. Further time-patterned analysis produces a 10-group solution that technically improves on the three-group solutions produced by both approaches. This 10-group solution is then organized on theoretical and empirical bases to identify a four-category typology that differs significantly from that produced by Kuhn and Culhane (1998).

Methods

Data

To analyze whether the time-patterned approach produces dissimilar results from the time-aggregated approach, this study tracks the study design of Kuhn and Culhane (1998). It does this because the Kuhn and Culhane design (1998) is the major empirical support for the prevailing three-category typology. Thus, the current study employs a recent version of the administrative data set these researchers used: the Single Client Information Management System (SCIMS) of the City of New York. For all single adults (age 18 or older) who enter New York City-supported shelters unaccompanied by mates or children, SCIMS records dates of each shelter entrance and exit and the information each person provides, at initial entrance, on his or her demographic characteristics, substance abuse, mental health, and physical health.

The study also follows the Kuhn and Culhane design (1998) in identifying the study population: people who first entered a New York City-supported shelter between January 1, 2000, and December 31, 2003. (Kuhn and Culhane select people first entering a shelter between January 1, 1988, and September 30, 1992.)³ As do Kuhn and Culhane in their study, the current analysis follows the shelter use of each person for 3 years from the date of his or her initial entrance. This results in 40,169 cases.⁴ To make the analysis more technically tractable, the current study randomly samples approximately one in eight of these cases, resulting in a sample of 5,000 people. This size is both large enough to discern groups with sufficient numbers of cases and yet small enough to be computationally workable for the software used in the analysis. The sample strongly represents the population of 40,169 cases from which it is drawn. A comparison of the sample with the population on demographic characteristics, shelter use measures, and theorized co-

variates shows that the largest difference in proportions is 0.007. No difference is statistically significant.⁵

Measurement

The time-patterned and time-aggregated analyses use the same temporal information but capture it differently. The current study's time-aggregated analysis follows Kuhn and Culhane (1998) in measuring shelter use frequency and duration: it counts each person's shelter stays and the total amount of time that he or she is sheltered over the 3-year observation period. Frequency is calculated by the number of shelter stays separated by at least 30 continuous days out of shelter. Duration is calculated by summing the number of days sheltered across all stays. Both calculations replicate the Kuhn and Culhane (1998) study.

By contrast, the time-patterned analysis breaks up the 3-year observation period into segments of 30 days (called a month). In each month, a person can be sheltered for between 0 and 30 days. The number of days sheltered each month is the input for the time-patterned analysis. The study uses 30-day time periods because using shorter periods (e.g., weeks) creates too many time periods for a tractable analysis; using longer periods (e.g., quarters) is analytically unnecessary and increases measurement error. Measuring time in 30-day periods is consistent with New York City's definition of the end of a shelter spell as more than 30 continuous days of absence from a shelter, and 30-day periods are consistent with Kuhn and Culhane's (1998) analysis. In addition, the current study uses the exact number of days sheltered in each time period because this approximates Kuhn and Culhane's use of the total number of days sheltered to measure duration.

In sum, the time-patterned analysis captures frequency and duration, as does Kuhn and Culhane's study (1998) but measures them by how they are distributed, not summed, over time. This measurement of frequency and duration follows from capturing the timing and sequence of events, features that Kuhn and Culhane do not capture. Combined, these measures allow the analysis to identify temporal patterns of homelessness in each person's life. In addition, for both the time-aggregated and time-patterned analyses, this study measures theorized covariates exactly as do Kuhn and Culhane.⁶

Analytic Techniques

To carry out the time-patterned analysis, this study uses optimal matching (OM; see the appendix for a description).⁷ In brief, OM first calculates a value that expresses how different the pattern of each individual's history is from the pattern of every other person in the data set. This value is obtained by calculating the smallest (or optimal) sum of the number of weighted changes needed to transform the homeless

sequence of one person into the exact same sequence of another. Doing this for all pairs of individuals in the data set produces an n -by- n dissimilarity matrix in which cell values reflect how different each person's sequence is from that of every other person. For people with relatively similar histories, this cell value will be small compared to the value for people with relatively different histories. The study uses the Transition Data Analysis software package to derive the dissimilarity matrix (Rohwer and Pötter 1999).

This matrix is then cluster analyzed or otherwise analyzed to group together people with relatively similar matrix values (i.e., people with relatively similar histories). These groups identify typological categories. Thus, OM generates typologies by combining cases; this contrasts with the usual way of forming typologies by combining values of variables. One can cross walk between these two ways of thinking by considering OM from the variable perspective. In this instance, cases that share sequences of values of the relevant phenomenon (e.g., homeless status) are grouped together. Thus, the unique combinations of values that define typological categories are found, in OM, not in combinations of particular values but in combinations of strings of values that are commonly shared across cases at each point in time, considering the sequence as a whole. (Of course, unique combinations are more approximated empirically than attained. For example, Bailey [1994] argues that unique combinations are only obtained in conceptual, not empirical, typologies.)

Furthermore, from the OM perspective, the dimensions being captured may not be obvious. For example, in matching people who stop and start homeless episodes at the same time, OM captures homeless duration (i.e., how much homelessness an individual has endured). It does so not by grouping individuals based on the number of nights they were homeless but by grouping them based on similarities in how long each episode of homelessness and nonhomelessness lasts, when the episodes occur in the sequence, and what comes before and what comes after each episode.

The present study generates these groups by using cluster analysis to analyze the dissimilarity matrix. Specifically, it uses Ward's method. This algorithm identifies clusters by finding solutions at each stage that produce the smallest change in the total within-group sum of squares, calculated as the within-group sum-of-square deviations from the mean of each cluster that are then summed across all clusters (Everitt, Landau, and Leese 2001).⁸ For example, having found a 10-group solution, Ward's method then finds the nine-group solution that has, among all possible nine-group solutions, the smallest increase from the 10-group solution in the total within-group sum of squares. This analysis uses Ward's method because its sum-of-squares criterion is convenient for discussing within-group homogeneity, and researchers have found that

it usefully clusters the OM dissimilarity matrix (see, e.g., Stovel and Bolan 2004; Stark and Vedres 2006). This study uses the Ward's method algorithm in ClustanGraphics 8.0 (Wishart 2004. For more technical descriptions of OM, see Sankoff and Kruskal [1983]; Abbott and Hrycak [1990]; for housing-relevant examples of its use, see Clark, Deurloo, and Dieleman [2003]; Stovel and Bolan [2004]; for sociological critiques of OM, see Levine [2000] and Wu [2000]).

The time-aggregated analysis in this study employs the method used by Kuhn and Culhane (1998). The number of days each person is sheltered and the number of his or her shelter episodes are simultaneously cluster analyzed. To make the time-patterned and time-aggregated analyses more comparable, this analysis uses Ward's method, rather than the nearest-centroid sorting technique used by Kuhn and Culhane (1998). They use this algorithm because it can cluster large data sets like theirs ($n = 73,263$). The much smaller data set in the current study removes this need. Using a different clustering technique is not important; this study's analysis of the effects of different ways of conceptualizing and measuring time-based concepts is unaffected by clustering technique. Although this difference is unimportant, this study did derive a three-group solution using nearest-centroid sorting. A comparison of these results to the time-aggregated three-group results found using Ward's method shows no statistically significant differences for group size, days sheltered, or average number of shelter episodes. It is important to note that, as the discussion previously observed, because the data used in this study follow that used by Kuhn and Culhane (1998), the resulting typologies characterize homeless shelter use and not a broader experience of homelessness.

Three-Group Analyses

Time-Patterned Analysis

Table 1 shows the three groups (A, B, and C) produced by the time-patterned analysis. To illustrate the kinds of cases in each group, the table reports 10 cases selected at random from each group. Numbers in cells are the number of days people were sheltered for each 30-day period; blank cells indicate that people were sheltered for zero days in that period.

The table suggests that group A resembles the theorized type of transitionally homeless. It is the largest group, including 74.5 percent of the sample, and people in this group are sheltered relatively briefly and have relatively little shelter experience after leaving. Group C resembles the theorized type for the chronically homeless. It is the smallest group, including 6.1 percent of all cases, and people in group C have relatively very long, continuous shelter stays and few nonshelter breaks. However,

Table 1

TIME-PATTERNED THREE-GROUP SOLUTION

CASE	30-DAY Time Periods																																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
Group A:																																						
A-1*	27	10																																				
A-2	1																																					
A-3	1																																					
A-4	2																																					
A-5	8																																					
A-6	2																																					
A-7	30	30	30	15																																		
A-8	20	9	30	4																																		
A-9	28	28	12	11	28	19	28	18																														
A-10	28	27	4					6	27	1																												
Group B:																																						
B-1	30	30	30	30	30	30	30	30	30	30	9																											
B-2	6	8	29	30	30	29	30	30	30	30	30	30	1																									
B-3	29	29	30	30	28	29	30	28	29	27	15	27	6																									
B-4	30	30	30	30	30	30	30	30	30	30	30	2																										
B-5	26	30	30	30	30	29	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	
B-6*	27	28	27	29	30	28	30	30	22	15	28	21	24	12	26	11	26	25	28	8	3	5																
B-7	30	30	30	30	26	30	25	30	29	30	30	30	30	30	30	25	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	
B-8	16	28	30	30	30	30	29	30	30	30	30	30	30	30	30	17	10																					
B-9	28	24	30	30	30	24	29	17																														
B-10	30	30	30	30	30	30	30	30	30	30	30	29	30	30	30	28	30	28	30	30	29	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	
Group C:																																						
C-1	30	30	30	7						7	27	30	20	1	6	4	22	19	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	
C-2	29	26	30	27	28	1					17	30	30	30	30	30	27	28	27	28	29	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	
C-3	30	30	27	30	29	28	30	29	30	30	30	20	28	26	29	30	27	28	27	28	29	28	29	28	29	28	29	28	29	28	29	28	29	28	29	28	29	28
C-4	30	30	30	30	30	30	30	30	30	30	30	27	30	29	27	29	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
C-5	28	30	29	29	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
C-6	27	25	27	25	28	24	28	28	28	28	30	29	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
C-7	30	30	30	29	30	30	29	21	13	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
C-8	30	10	30	30	30	30	6			1	30	30	29	30	16	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	
C-9	29	30	30	30	30	30	30	30	30	30	30	30	30	2																								
C-10*	22	29	30	30	30	30	26	30	19	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	

NOTE.—Numbers in cells are the number of days people were sheltered for each 30-day period; blank cells indicate that people were sheltered for zero days in that period.
* This case is an example of the population of illustrative cases in that group (A, B, or C).

Table 2

THEORETICAL AND EMPIRICAL AVERAGE TOTAL WITHIN-GROUP DISTANCES FOR
THE THREE-GROUP TIME-PATTERNED ANALYSIS

TYPOLOGY CATEGORY TYPE	THEORETICAL IDEAL-TYPE EXPECTATIONS			EMPIRICAL ILLUSTRATIVE CASE FINDINGS		
	Transitional	Episodic	Chronic	Group A	Group B	Group C
Transitional	.0	> .0	> .0	7.3	50.2	105.2
Episodic	> .0	.0	> .0	51.7	45.4	70.8
Chronic	> .0	> .0	.0	120.8	79.0	25.6

group C also contains people with strong in- and out-of-shelter histories. This leaves group B, then, to represent the theorized type for the episodically homeless. Comprising 19.6 percent of all cases, group B is small relative to group A and does include people with histories of moving in and out of shelter, as the episodic category was theorized. But durations for these people are longer than expected, and the group is dominated by people sheltered for a relatively long, maximal (29–30 days of a time period), continuous (12–18 months) time, and then not at all. This evidence is weak support, at best, for the theorized episodic type.⁹

Another way to see how these results relate to theorized types is to identify people in each group with shelter and nonshelter histories that are akin to what is expected for the group and then to calculate how different their histories are from all other histories in the group. For any group, the expected histories should more closely match all histories than the histories not expected for that group. For example, histories in group A that typify the transitional type should be less different from all histories in group A than histories that typify groups B and C. The theoretical expectations panel of table 2 shows the logic of this reasoning. For each theorized ideal type, histories are all exactly the same, are found only in their theorized category, and are the only cases in that category (0.0 dissimilarity on the diagonal). The idealized histories from the other groups differ from these histories (> 0.0 dissimilarity in the off diagonal).

The empirical analysis identifies the people in each group who have histories consistent with the theorized expectations for that group. The analysis identifies these people by randomly sampling each group's population of illustrative histories and by constructing dissimilarity matrices for each group using the dissimilarities initially produced by the OM analysis. (Histories marked with an asterisk in table 1 are examples of the population of illustrative histories in that group.)¹⁰ The analysis then calculates how different, on average, each sample of illustrative histories is from all other histories in each group.¹¹

Table 3

TIME-AGGREGATED THREE-GROUP SOLUTION

MEASURE	GROUP		
	A (Transitional)	B (Episodic)	C (Chronic)
% of <i>N</i>	73.4	4.6	22.0
Average no. episodes	1.3	4.7	1.5
Average no. days sheltered	44.7	243.5	509.1

NOTE.—*N* = 5,000 cases.

Results in the empirical findings panel of table 2 suggest that the average total within-group difference in group A is smallest for histories illustrative of the transitional type (7.3, as opposed to 51.7 for episodic and 120.8 for chronic). In group B, this difference is smallest for histories illustrating the episodic type (45.4 as opposed to 50.2 for transitional and 79.0 for chronic). In group C, histories that illustrate the chronic type show the smallest difference (25.6, as opposed to 70.8 for episodic and 105.2 for transitional). These results support this study’s previous characterizations of each group. As before, support for group B is much weaker than support for the other groups. Group B has the lowest variation among differences across illustrative histories. This indicates that its illustrative histories are almost as similar to all other histories in group B as are the chronic and transitional illustrative histories. In addition, all histories illustrating the transitional type are found in group A, and all histories illustrating the chronic type are found in group C, but histories illustrating the episodic type are found in all three groups.

These analyses suggest that some support exists for the three-category hypothesis but that support is not strong. The next issue to address, then, is how well does this time-patterned approach perform relative to a time-aggregated approach? And how well do both generate relatively homogeneous groups? The ability to do so is crucial in forming a typology.

Time-Aggregated Comparison

To compare these results with those from a time-aggregated analysis, the authors reproduce Kuhn and Culhane’s (1998) three-group study using the current data set. Table 3 reports the results for the measures of frequency (average number of episodes) and duration (average number of days sheltered). The results are strongly consistent with expectations for the transitional, episodic, and chronic categories. They also are consistent with the group characteristics that Kuhn and Culhane initially find.¹² The results presented in table 3 are compared with results

Table 4

HOMOGENEITY OF THREE-GROUP FINDINGS FOR TIME-PATTERNED
AND TIME-AGGREGATED ANALYSES

	WITHIN-GROUP SUM OF SQUARES		
	Raw Total SS	Avg SS	Ratio of Group Avg to Total Avg
Time-patterned analysis:			
Transitional	38,576	10.4	.33
Episodic	30,756	31.6	1.01
Chronic	8,792	28.8	.92
Total within three-group SS	78,123		
Total data set (one-group) SS	156,502	31.3	
Proportion explained of total SS		.50	
Time-aggregated analysis:			
Transitional	751	.2	.20
Episodic	231	1.0	1.00
Chronic	897	.8	.82
Total within three-group SS	1,879		
Total data set (one-group) SS	4,999	1.0	
Proportion explained of total SS		.62	

NOTE.—Avg = average; SS = within-group sum of squares. The difference in raw total sum of squares is because the opportunity for variation is greater in the time-patterned analysis than in the time-aggregated analysis. The time-patterned analysis measures duration and frequency over 36 time periods, but the time-aggregated analysis effectively collapses these measures into one moment. Thus, people have more opportunity to differ on these measures, and, in this data set, they do.

from the time-patterned analysis for within-group homogeneity and for covariate associations. The analytic usefulness of typological concepts increases with the homogeneity of their empirical referent and with how well they are associated with measures of other concepts.

The analysis assesses within-group homogeneity by calculating total within-group sum-of-squared deviations from the group average (cluster centroid). This measure is based on the principle used in optimization cluster analysis to assess cluster homogeneity (Everitt et al. 2001). Also, the clustering algorithm used in this study groups people by minimizing changes in the total within-group sum of squares. Table 4 shows the raw within-group sum of squares for each group for each approach. To standardize for differences in group size, the table also shows the average within-group sum of squares.

As table 4 reports, the average sum of squares in the time-patterned results are 10.4 for the transitional group, 31.6 for the episodic group, and 28.8 for the chronic group. The average is 31.3 for the entire data set.¹³ As the ratios of these measures show (third column in table), the heterogeneity in the episodic and chronic groups is about as great as that in the full data set (i.e., if the data set is treated as one cluster). The transitional group, however, is markedly more homogeneous. These results are consistent with the degree of variability found in the initial

time-patterned analysis. In addition, the time-patterned solution explains about 50 percent of its total data set variability, and the time-aggregated solution explains about 62 percent of its total data set variability; both leave a good deal to be explained.¹⁴ Further, this analysis suggests that neither approach generates a homogeneous chronic group or a homogenous episodic group.

To assess the relations of these groups with measures of other concepts, analyses examine the groups' associations with covariates theorized by Kuhn and Culhane (1998) and measured here: age, medical problems, mental health problems, and substance abuse problems. The analyses consist of comparing two groups at a time on each of these variables. Measurement of the covariates follows Kuhn and Culhane: age is trichotomized at 30 and 50 years, and the other covariates are measured in SCIMS as whether or not people have a particular kind of problem. Table 5 reports the results of these analyses. It shows the directions of theorized relations, odds ratios, and associated 99 percent confidence intervals, as well as statistical significance.¹⁵ For example, in the first row of results from the time-patterned analysis, the table reports that the transitionally homeless group is expected to include more people under age 30 than the chronically homeless, that the odds ratio is estimated to be a statistically significant 2.52, and that this level supports the theory.

Results from the time-patterned analysis generally suggest that the transitional-chronic and transitional-episodic comparisons show strong support for the claims of the theory. All relations with all covariates are in the right direction and are statistically significant (save that for mental health problems; it is statistically significant with a 95-percent confidence interval).¹⁶ However, except for age, odds ratios for chronic-episodic relations are not statistically significant. Also, the odds ratios for mental health and substance abuse problems are in the wrong direction. In the time-patterned analysis, the chronic-episodic odds ratios are stronger, and more of the associations are statistically significant at .01 than are those in the time-aggregated results. Only the statistical significance for mental health problems in the transitional-chronic comparison and the estimated odds ratio for substance abuse problems in the transitional-episodic comparison are less robust for the time-patterned groups than for the time-aggregated groups. In both analyses, the chronic-episodic relations are, for the most part, not statistically significant. Note, however, that the confidence intervals for statistically significant odds ratios in one analysis overlap with intervals for corresponding odds ratios in the other analysis. This overlap makes it impossible to say that the odds ratios are statistically distinguishable. All one can say is that the odds ratios in the time-patterned analysis are, in general, more likely to be stronger, in the right direction, and statistically significant than the comparable odds ratios in the time-aggregated analysis.

Table 5
THEORIZED AND ESTIMATED COVARIATE RELATIONS FOR TIME-PATTERNED AND TIME-AGGREGATED APPROACHES

	TRANSITIONAL-CHRONIC			TRANSITIONAL-EPISODIC			CHRONIC-EPISODIC				
	Theory	OR	99% CI	Theory	OR	99% CI	Theory	OR	99% CI		
Time-patterned analysis:											
Age < 30	> 1.0	2.52	1.48	4.28	1.38	1.07	1.79	< 1.0	.55	.31	.97
Age > 50	< 1.0	.38	.28	.51	.59	.49	.73	> 1.0	1.58	1.12	2.23
Mental health problems	< 1.0	.71 ^a	.47	1.07	.74	.58	.95	< 1.0	1.04 ^a	.67	1.62
Medical problems	< 1.0	.50	.34	.73	.67	.53	.85	> 1.0	1.34 ^a	.88	2.03
Substance abuse problems	< 1.0	.65	.43	.99	.69	.54	.88	< 1.0	1.06 ^a	.67	1.68
Any problem	< 1.0	.44	.28	.69	.58	.45	.74	< 1.0	1.32 ^a	.81	2.18
All three problems	< 1.0	.19	.08	.46	.39	.21	.74	< 1.0	2.03 ^a	.79	5.25
Time-aggregated analysis:											
Age < 30	> 1.0	1.69	1.31	2.18	1.34 ^a	.83	2.17	< 1.0	.80 ^a	.47	1.34
Age > 50	< 1.0	.46	.38	.56	1.13 ^a	.74	1.74	> 1.0	2.44	1.56	3.82
Mental health problems	< 1.0	.72	.56	.91	.64 ^a	.41	1.00	< 1.0	.90 ^a	.56	1.44
Medical problems	< 1.0	.54	.43	.67	1.06 ^a	.66	1.70	> 1.0	1.98	1.21	3.24
Substance abuse problems	< 1.0	.68	.53	.86	.61	.39	.94	< 1.0	.90 ^a	.56	1.44
Any problem	< 1.0	.49	.38	.62	.70 ^a	.44	1.10	< 1.0	1.44 ^a	.88	2.36
All three problems	< 1.0	.26	.15	.47	.41 ^a	.14	1.24	< 1.0	1.57 ^a	.50	4.87

NOTE.—OR = odds ratio; CI = confidence interval. The Theory columns indicate theorized direction of odds ratio. Empty cells in the Theory columns indicate that there is no applicable theory or that the theory is ambiguous.

^a Odds ratio is not statistically significant at $p < .01$.

Conclusions for Three-Group Analyses

These analyses suggest several things. First, the time-patterned and time-aggregated approaches identify similar three-category typologies, and these typologies weakly support the theorized categories. Second, the within-group homogeneity has the same character in both analyses (if one compares across groups within each analysis). Third, the time-patterned results are more strongly and statistically significantly related to theorized covariates than those from the time-aggregated analysis. Fourth, and most important, the three-group solutions identified by both approaches create groups that retain a good deal of heterogeneity. This is evidenced by the proportion of remaining data set variation and by the absence of support for theorized relations concerning the episodic and chronic groups. In both analyses, the transitional group performs well; it is relatively homogeneous and has strong, statistically significant relations with covariates. By contrast, in both analyses, the chronic and especially the episodic groups perform badly on both homogeneity and strength of relations. These findings suggest the utility of further analysis to gain a stronger understanding of the temporal structure of shelter use.

The weakness of the theorizing cited by Kuhn and Culhane (1998) also suggests further analysis.¹⁷ This theorizing argues that the transitional and chronic categories have the same frequency (low) but vary on duration. All other homelessness is therefore episodic; that is, homelessness consists of a short-lived emergency, a long-term condition, or something else. This theory may oversimplify short-term and long-term homelessness; it also makes episodic homelessness a residual category. The analysis presented here supports this characterization of the episodic category as residual and supports the notion that long-term homelessness is oversimplified.

Extant theory and the empirical research based on it have been an important starting point for understanding the temporal structure of homelessness. The next section builds on this start by identifying a time-patterned solution that addresses the issues just discussed. This solution provides an alternative to the three-category typology, an alternative that has greater homogeneity and a different characterization of time-based homelessness than that found by using a time-aggregated approach.

Time-Patterned Analysis

To identify potential time-patterned solutions, the following analysis uses information from the previous time-patterned analysis and two guides. The first guide is statistical and is presented in figure 2. The left vertical axis presents the scale for total within-group sum of squares. The right axis presents the scale for the *t*-statistic, which evaluates the statistical

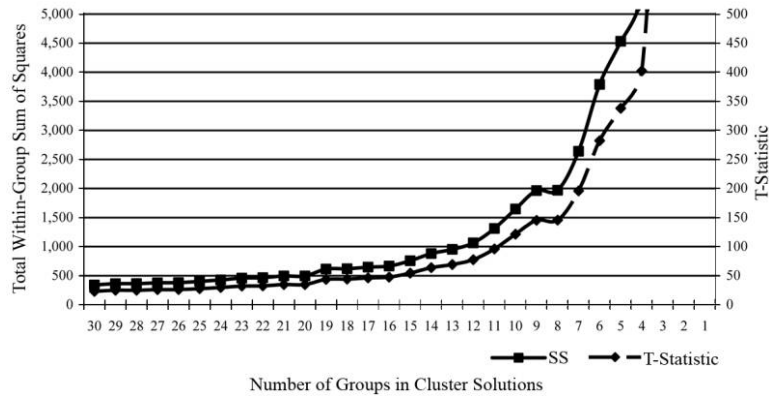


FIG. 2.—Total within-group sum of squares (SS) and t -statistic for fusion values for cluster solutions 30 through 1.

significance of the fusion values for each solution. Favored solutions are those prior to a relatively large change in statistically significant fusion values; that is, solutions are excluded at the point where reducing the number of groups causes a relatively large increase in the total within-group sum of squares (Wishart 2005). By this criterion, the figure suggests that the nine- to 12-group solutions may comprise one possibly useful set.

The second guide is to rely on theories about temporal homelessness when those theories are relatively strong. One such theory concerns temporary homelessness, a version of which is articulated in Kuhn and Culhane's (1998) transitional category and by other researchers (e.g., Rossi, Fisher, and Willis 1986; Culhane and Metraux 1999; Burt et al. 2001). Research suggests that temporary homelessness is comprised of a relatively large population of people who become homeless after one-off events (e.g., a fire that razes a house, a unique financial moment, a breakup in an intimate relationship, a transition from institutionalized care to a home of one's own). In the context of shelter use, these people enter shelters once, for a very short time, and they never return. Thus, from among the nine- to 12-group solutions, the authors choose solutions in which people were sheltered only once and for no more than 30 days, and, from among those, the solution that is most parsimonious (i.e., has the fewest groups). The 10-group solution meets these criteria.¹⁸

Results

Table 6 reports exemplars for the 10-group solution. These are cases in each group that have the smallest within-group average dissimilarity to all other cases in the group (Wishart 2004) and that appear to represent the dominant temporal structure of cases in the group. For ease of

Table 6
EXEMPLARS FOR THE TIME-PATTERNED 10-GROUP SOLUTION

GROUP	% OF N	30-DAY TIME PERIODS																																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
Temporary: A	37.2	3																																				
Structured- continuous:																																						
B	18.8	28	8																																			
C	12.1	29	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
D	6.6	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
E	6.5	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
F	2.8	30	29	30	30	30	30	29	30	30	30	30	30	30	30	30	30	30	30	30	30	29	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
G	3.3	29	30	30	30	29	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Subtotal	50.2																																					
Structured- intermittent:																																						
H	2.3	29	29	29	30	6																																
I	4.1	26	30	30	7				20	30	30	29	30	30	12																							
Subtotal	6.3																																					
Unstructured- intermittent:																																						
J	6.3																																					

NOTE.—There are too many different kinds of the unstructured-intermittent pattern to identify a meaningful exemplar for group J. Numbers in cells are the number of days people were sheltered for each 30-day period; blank cells indicate that people were sheltered for zero days in that period.

exposition, the 10 groups were organized into four sets based on theoretical expectations for the temporary category and on observed similarities across subsets of groups in the pattern of shelter and nonshelter stays. The four sets are temporary, structured-continuous, structured-intermittent, and unstructured-intermittent. (Weisburd et al. 2004 use a similar strategy to organize and describe trajectories found through group-based modeling techniques.)

The first set fits the theorized temporary homeless. It refers to people who have one very brief shelter stay in the initial 30-day period and no shelter reentrance. This finding is not surprising, because the selection of a solution is made from available solutions that best represent this temporary type. Theory and empirical evidence from New York City shelters lead the authors to expect such a pattern, but the time-patterned approach and OM offered no assurance that a pattern fitting the temporary category would be found in any solution. Finding such a pattern encourages the authors' confidence in this particular solution, though there are also other solutions with such a temporary pattern.

The structured-continuous set involves one continuous shelter stay for the maximum time each 30-day period and little, but some, reentrance. Major group variation in this set arises in the duration of the stay. Although people in group G are still sheltered at the end of the observation period, this group shares the set traits of maximal and continuous time in shelter. A longer observation period would likely have shown that additional groups have the same pattern and leave at some point after 36 months. Not shown in the exemplar is that the groups contain people who do return, but their return stay is generally for less than 30 days and for one 30-day period.

The structured-intermittent set includes groups with sequences of shelter and nonshelter use. During periods of shelter use, the stay tends to be for most or all of the 30-day period, and periods of use occur at different points in the observation period. The two groups in this set show variations in timing and duration. Both have a short shelter sequence after people enter, but group H has another sequence almost 18 months later, and the sequence lasts for a year. Group I has a second sequence 6 months after the first one, and the second sequence lasts for 6 months. These histories suggest that group H may be more likely to return to shelter after the observation period. Both groups contain variability not displayed in the exemplars; in particular, people in both groups reenter for brief periods of time and stay for less than the maximum time.

The unstructured-intermittent set contains highly varying sequences of shelter and nonshelter use. Sequences in this set are more intermittent and more unstructured than those for other sets. Also relative to these sets, shelter use is highly variable in the amount of time sheltered during each 30-day period. Because histories are so variable across in-

dividuals, identifying an exemplar would not be as informative as it is for the other groups. To graphically indicate this group, appendix table A1 shows five randomly drawn cases.

Evaluating Results

This analysis is undertaken to improve the homogeneity and reduce the substantive ambiguity of the time-patterned three-group solution. Also, to further demonstrate the utility of the time-patterned approach itself, the analysis is expected to produce different substantive results from those produced using the time-aggregated approach. The 10-group solution is therefore first compared with this study's time-patterned three-group solution and then with comparable solutions generated by the time-aggregated approach.

For the time-patterned three-group and 10-group solutions, within-group homogeneity is measured by calculating weighted means for the within-group sum of squares of the groups comprising each solution. This average sum of squares is 9.1 for the 10-group solution compared to 15.6 for the three-group solution, a 42 percent reduction in within-group sum of squares. To measure total data set homogeneity, the analysis calculates an average total sum of squares and estimates how much of this variation each solution explains. The 10-group solution explains 71 percent; the three-group solution explains 50.1, also an improvement of 42 percent.

The 10-group solution is substantively different from the three-group solution. Perhaps one similar category, the transitional, or temporary, is found in both analyses. A segment of the chronically homeless is now seen to be one group among many others comprising the structured-continuous set. The episodic and chronic are now seen to comprise two very different sets: structured intermittent and unstructured intermittent. Thus, the 10-group solution produces types that are very different from those produced by the three-group solution.

These types are also different from those produced by two comparable time-aggregated solutions: a technically optimal time-aggregated solution and a time-aggregated 10-group solution. (Results not shown but available on request.) To identify an optimal time-aggregated solution, the study uses the same criteria employed to identify an optimal time-patterned 10-group solution. These criteria identify the point at which further clustering overly increases the total within-group sum of squares relative to the number of groups in the solution. The result of this optimizing analysis is a six-group solution. A second analysis identifies a time-aggregated 10-group solution in order to compare the two approaches across solutions with the same numbers of groups.

The result of both analyses is a more refined version of the time-aggregated three-group solution. The additional groups in both solu-

tions are simply slightly different variations on the themes of the transitional, episodic, and chronic categories.¹⁹ These analyses do not yield typologies that are different from the time-aggregated three-group typology nor a typology that is the same as that produced by the time-patterned analysis. Thus, the time-patterned approach generates a typology that could not have been identified using a time-aggregated approach.

Discussion

Michael Sosin's (2003) discussion of theories on the causes of homelessness points out that a complete explanation for homelessness has to account for why homeless people move between different housing circumstances. He also points out that such an explanation has to include the effects on those transitions of factors beyond individual characteristics. Using a time-patterned approach to form typologies creates opportunities to generate and test such explanations that time-aggregated analysis cannot. Time-patterned categories identify similar transitions from one status to another and at similar specific moments in time. An explanation of the nature and timing of these transitions must incorporate factors different from (but in addition to) individual traits, because how traits play a role in effecting these changes depends on structural conditions. They cannot by themselves generate changes in sequences.

Structural conditions provide explanations in two ways. One is that these conditions change. For example, new policies governing shelters go into effect, or changes are made in how mental health services are offered, or the criminal justice system changes how it responds to street homelessness. Another is that how structural conditions operate generates temporal patterning. Homeless people with mental health problems, for example, are subject to the procedures, practices, and organization of the mental health system. They also may be subject to features of the criminal justice and health care systems. When such people initially use shelters, how long they stay and when they return are likely determined by their individual exposure to and responses to these systems. To explain the structured-continuous groups, for example, it is not enough to say that they are linearly related to age (as they are). The explanation also must specify mechanisms (structural conditions) that translate age into the observed groups.²⁰

The remainder of this section illustrates this argument by imagining theoretically suggestive explanations for the temporary and structured-continuous sets.²¹ These are not the only possible explanations and are not tested here. They are intended only to suggest the utility of the time-patterned approach for constructing a causal theory of homelessness.

Temporary

People in this set leave shelters within 30 days, never to return. The theoretically relevant questions are why they remain sheltered so briefly and why they never return (at least during the observation period). Not returning is key because, in this study's groupings, returning would place a person in a different set.

One set of possible answers has to do with the effects of shelter environment (e.g., their rules, regulations, physical condition, and social circumstances). For example, shelters may repel people who have low tolerance for rules or fear living among strangers.²² Such structural conditions select against long stays for certain kinds of people. Further, if relevant individual traits and if key shelter features (organization, physical conditions, or social circumstances) do not change over the observation period, the person is unlikely to reenter the shelter. Of course, whether people reenter also depends on their ability to negotiate structural conditions outside the shelter.

These suggestions are consistent with the little data available. Relative to other groups, the temporary group is comprised of substantially more white people, and they have relatively low levels of mental health, physical health, and substance abuse problems. These qualities are likely to help individuals negotiate living conditions outside shelters and so allow them to eschew the problematic conditions of shelter life.²³

Structured-Continuous

People in this set stay continuously sheltered each month until they leave, and groups are determined by successively later time points of leaving. (If graphed, these time points would resemble a strongly monotonic, concave curve.) This suggests that these individuals find shelter conditions are more or less acceptable and can or want to stay fully sheltered while using them. Also, people in this category may only return to shelter once, sometimes twice, for a short time, suggesting that they are pretty much finished with using shelters by the time they leave. To explain this patterning, one might imagine that experience with institutionalized living (e.g., in jail, prison, hospitals, as well as shelters) enables these individuals to deal with shelter conditions and that, as people have more problems, they have greater need for services that shelters provide. One also might imagine why people leave so discretely: time-based policies generate leaving or provide opportunities for doing so. For example, individuals meet formal shelter-time criteria for receiving permanent housing. In addition, shelter conditions may make people's problems more difficult, even as those problems make shelter living necessary. Perhaps as a result of individual traits, people deal with this tension for variable lengths of time.

Available data allow only a small and very tentative examination of this reasoning concerning the structured-continuous set. The argument suggests that the successively greater durations of the six groups comprising this category are positively and monotonically related to age (as a measure of experience), mental health problems, physical health problems, and substance abuse problems. The authors find these relationships.

The authors reiterate that these are some arguments that can be imagined to explain the time-patterned results but that the current data set does not have evidence to rigorously test these explanations. Rather, the authors' point has been to provide examples of explanations to illustrate how time-patterned analysis constructs the lives of homeless people in ways that enable research to imagine how structural conditions interact with individual traits to generate histories of homelessness. More definitive typologies than the ones presented here would identify distinct temporal conceptualizations of homelessness that have distinct sets of causes and effects.

Conclusion

Constructing Time-Based Typologies

The major point of this analysis is to suggest the utility of forming temporally based homeless typologies by incorporating more temporal information than is commonly used. The dominant typology in homeless research and policy making uses much less temporal information than is available. This study adduces evidence concerning group homogeneity and the substantive nature of the identified groups, suggesting that it is possible to construct groups that, compared to those specified in previous research, are more homogeneous and reveal a different picture of the temporal nature of shelter use in people's lives. This study does not argue, however, that a time-patterned approach is the only right way to typologize homelessness. The approach should be determined by theory and the uses to which the typology is put. If theory necessitates that time be aggregated (e.g., if the theory argues that only the amount of time homeless matters), then it is not productive to test that theory by unpacking the aggregation. If a theory argues, however, that the sequence of housing events explains the condition of individuals' mental health or that structural conditions explain people's histories of homelessness, then unpacking how events occur over time will be useful. Further, as this study shows, unpacking the temporal aspects of events can help develop concepts not theoretically anticipated.

Number of Groups

There can be more than one correct number of groups for forming a typology.²⁴ With either the time-patterned or the time-aggregated ap-

proach, it is possible to identify typologies that have different numbers of technically valid groups. Kuhn and Culhane (1998), for example, do not discount that more or fewer than three groups could be found. The analytic issue is how to calibrate parsimony in relation to accuracy, particularly since different numbers of technically valid groups can be correct even with the same data set. Specifically, in the current analysis, the issue is how to trade off a smaller number of temporally heterogeneous groups for a larger number of temporally homogeneous groups. As this study just argued regarding the choice of analytic approaches, the theory and the uses to which the analysis is put determines the number of groups chosen. Theorizing three categories and finding three empirical groups to support that theory may be sufficient to explain some phenomena or to improve policy making, but this does not mean that findings cannot be improved technically and advanced conceptually, as this study shows. This argument is true for different solutions that are technically valid, but it can also be true, as Andrew Gelman and Ronald Rubin (1995) show, even when a possible solution is not technically valid.

Structural Theorizing

Last, this study emphasizes the utility of a time-patterned approach for structural theorizing. By identifying previously unknown housing durations and transitions shared by groups of people, this study can better theorize the role of specific structural factors in generating these sequences. For example, the research allows theory to account for how transitions are generated and how a new status is sustained by such factors as the organization of and access to services, shelters, the health care system, and the criminal justice system, as well as by changes to all these. Structural theorizing is of course possible with a time-aggregated approach, but it does not create the opportunity to use structural conditions to explain change. As a result of using a time-patterned approach, shared histories of movements between housed, homeless and other living situations, as well as the amount of time in each situation, can be analyzed by considering how structural conditions and individual traits interact to produce observed histories. Thus, a time-patterned approach allows researchers to generate more complete explanations for the causes and effects of homelessness.

Appendix A

Optimal Matching Analysis

Optimal matching has two steps. The first step is to calculate an n -by- n dissimilarity matrix that expresses how different each sequence is from every other sequence. The second step is to analyze the dissimilarity

matrix in order to group similar sequences. Critical to the first step is setting weights (which are also called costs) to value the different transformations used to turn one sequence into another. This transformation happens by substituting, inserting, or deleting values into a sequence. Substitutions replace the value in one sequence with the same value from the sequence with which it is being compared; insertions and deletions (hereafter, INDELS) insert or delete values from a sequence relative to the values in the sequence with which it is being compared. Different substitutions and INDELS can be differently weighted. All possible transformations may not be equally important. In the current study, each person's history consists of counts of the number of days homeless over each consecutive 30 day period (i.e., the count ranges from zero to 30; over the observation period, there are 36 such time periods). Thus, it is necessary to assign weights for substituting, inserting, or deleting these values for one another (e.g., the weight for substituting 5 days sheltered in one sequence for 10 days sheltered in another).

Substitution Weights: 1–30 Shelter Days

The analysis utilizes the metric in the continuous measure of number of days sheltered in each time period to identify substitution values by simple subtraction (i.e., the weight of substituting 10 days sheltered for 7 days sheltered is 3). Evidence from the current data suggests, however, that the probability of leaving shelter in any 1 month declines quickly after the first few days and then stays more or less constant but may increase slightly near the end of the month. This suggests it may be more difficult for individuals to stay sheltered a second or third day than for a nineteenth or twentieth day. Substitution weights should reflect this behavior. Thus, to generate these weights for states 1–30, the study uses substitution weights that are the differences of the natural log of each number of days being compared.

Substitution Weights: Out of Shelter to in Shelter

Evidence from the current data suggests that entering a shelter (or reentering one) is much more difficult than remaining sheltered. To remain sheltered each night is to return to a place where rules, social character, and physical qualities are known, if not familiar. A person who remains sheltered has more or less successfully lived under these conditions. People entering anew, and even those reentering shelter, may be more likely to be concerned about these issues. Thus, the weight for substituting any number of days sheltered for days out of shelter (i.e., zero days sheltered) should be meaningfully larger than the weight for substituting any number of days sheltered for 1 day sheltered. For example, substituting 1 day sheltered for an out-of-shelter day should be meaningfully greater than substituting 2 days sheltered for 1 day sheltered. To accomplish this, the analysis arbitrarily assigns the value of .025 to out-of-shelter days, since subtracting the natural log of this

number from the log of any number of days in shelter provides a sufficiently large cost. Doing so sets the weight for substituting 1 day sheltered for 1 day out-of-shelter at just over 5 times the cost of substituting 2 days sheltered for 1 day sheltered.

INDEL Weights

These weights are typically set in relation to substitution weights. Their value depends on how much the analysis wants to value INDELS in forming the dissimilarity matrix. Employing INDELS emphasizes the importance of strings of similar values in a sequence (duration). Doing so deemphasizes the importance of when those values occur (timing). Technically, the greater the value of INDELS relative to substitution weights, the less likely is the algorithm to use them. (For discussions of the logic and meaning of INDELS, see Abbott and Hrycak 1990; Abbott and Tsay 2000; Lesnard 2006; Lesnard n.d.)

When sequence lengths are unequal, INDELS can be particularly useful; deleting and inserting values is equivalent to deleting and inserting time periods. Thus, deleting and inserting values equalizes sequence lengths. In the current study, because all lengths are equal, INDELS can be thought of as making the substitution weight matrix more subtle than it would be without INDELS: because sequence lengths are equal, every insertion requires a subsequent deletion (and vice versa) for each pair of sequence lengths to remain equal. Over the entire sequence, this is effectively a substitution operation.

The sequences in this data set are dominated by out-of-shelter time periods (i.e., zero days sheltered). This suggests the importance of setting INDELS relative to the weights for substituting being out of shelter for being sheltered for any number of days. The most useful results stem from setting the INDEL at 1.8. (This is .50 of the weight for substituting zero days sheltered for 1 day sheltered and is .26 of the maximum substitution weight.) With this value, INDELS can play a role in equating a time period by substituting out-of-shelter days for any number of days sheltered (or vice versa); they play no role in equating a time period by substituting any number of days sheltered for each other. In this latter instance, only substitution weights are used.

Table A1
RANDOM SELECTION OF CASES FROM UNSTRUCTURED-INTERMITTENT PATTERN

CASE	30-DAY TIME PERIODS																																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
J-1	15	4	1		5	13																														23	6
J-2	2																1	28	28																		
J-3	9										3	1																									
J-4	18																												4	17	12	1					
J-5	5						1							11	30	7																					

NOTE.—Numbers in cells are the number of days people were sheltered for each 30-day period; blank cells indicate that people were sheltered for zero days in that period.

References

- Abbott, Andrew. 2001. *Time Matters: On Theory and Method*. Chicago: University of Chicago Press.
- Abbott, Andrew, and Alexandra Hrycak. 1990. "Measuring Resemblance in Sequence Analysis: An Optimal Matching Analysis of Musicians' Careers." *American Journal of Sociology* 96 (1): 144–85.
- Abbott, Andrew, and Angela Tsay. 2000. "Sequence Analysis and Optimal Matching Methods in Sociology." *Sociological Methods and Research* 29 (1): 3–33.
- Bailey, Kenneth D. 1994. *Typologies and Taxonomies: An Introduction to Classification Techniques*. Thousand Oaks, CA: Sage.
- Bauer, Daniel J., and Patrick J. Curran. 2003. "Distributional Assumptions of Growth Mixture Models: Implications for Overextraction of Latent Trajectory Classes." *Psychological Methods* 8 (3): 338–63.
- Burt, Martha, Laudan Y. Aron, and Edgar Lee, with Jesse Valente. 2001. *Helping America's Homeless: Emergency Shelter or Affordable Housing?* Washington, DC: Urban Institute Press.
- Caton, Carol L. M., Boanerges Dominguez, Bella Schanzer, Deborah S. Hasin, Patrick E. Shrout, Alan Felix, Hunter McQuiston, Lewis A. Opler, and Eustace Hsu. 2005. "Risk Factors for Long-Term Homelessness: Findings from a Longitudinal Study of First-Time Homeless Single Adults." *American Journal of Public Health* 95 (10): 1753–59.
- Clark, William A. V., Marinus C. Deurloo, and Frans M. Dieleman. 2003. "Housing Careers in the United States, 1968–1993: Modelling the Sequencing of Housing States." *Urban Studies* 40 (1): 143–60.
- Culhane, Dennis P., and Stephen Metraux. 1999. "One-Year Rates of Public Shelter Utilization by Race/Ethnicity, Age, Sex and Poverty Status for New York City (1990 and 1995) and Philadelphia (1995)." *Population Research and Policy Review* 18 (3): 219–36.
- Culhane, Dennis P., Stephen Metraux, Jung Min Park, Maryanne Schretzman, and Jesse Valente. 2007. "Testing a Typology of Family Homelessness Based on Patterns of Public Shelter Utilization in Four U.S. Jurisdictions: Implications for Policy and Program Planning." *Housing Policy Debate* 18 (1): 1–28.
- Cunningham, Mary. 2009. "Preventing and Ending Homelessness—Next Steps." Metropolitan Housing and Community Center Newsletter, February. Urban Institute, Washington, DC.
- Doty, D. Harold, and William H. Glick. 1994. "Typologies as a Unique Form of Theory Building: Toward Improved Understanding and Modeling." *Academy of Management Review* 19 (2): 230–51.
- Eggleston, Elaine P., John H. Laub, and Robert J. Sampson. 2004. "Methodological Sensitivities to Latent Class Analysis of Long-Term Criminal Trajectories." *Journal of Quantitative Criminology* 20 (1): 1–26.
- Everitt, Brian S., Sabine Landau, and Morven Leese. 2001. *Cluster Analysis*. 4th ed. London: Arnold.
- Gelman, Andrew, and Donald B. Rubin. 1995. "Avoiding Model Selection in Bayesian Social Research." *Sociological Methodology* 25:165–73.
- Goering, Paula, George Tolomiczenko, Tess Sheldon, Katherine Boydell, and Donald Wasylenki. 2002. "Characteristics of Persons Who Are Homeless for the First Time." *Psychiatric Services* 53 (11): 1472–74.
- Jahiel, Rene, and Thomas F. Babor. 2007. "Toward a Typology of Homeless Families: Conceptual and Methodological Issues." Appendix B1 in *Characteristics and Dynamics of Homeless Families with Children: Final Report to the Office of the Assistant Secretary for Planning and Evaluation, Office of Human Services Policy, U.S. Department of Health and Human Services*, edited by Debra J. Rog, C. Scott Holupka, and Lisa C. Patton. Washington, DC: U.S. Department of Health and Human Services. <http://aspe.hhs.gov/hsp/homelessness/improving-data08/apb.htm>.
- Kertesz, Stefan G., Mary Jo Larson, Nicholas J. Horton, Michael Winter, Richard Saitz, and Jeffrey H. Samet. 2005. "Homeless Chronicity and Health-Related Quality of Life Trajectories among Adults with Addictions." *Medical Care* 43 (6): 574–85.
- Kuhn, Randall, and Dennis P. Culhane. 1998. "Applying Cluster Analysis to Test a Typology of Homelessness by Pattern of Shelter Utilization: Results from the Analysis of Administrative Data." *American Journal of Community Psychology* 26 (2): 207–32.

- Lesnard, Laurent. 2006. "Optimal Matching and Social Sciences." Working Paper no. 2006-01. Centre de Recherche en Economie et Statistique, Institut National de la Statistique et des Etudes Economiques, Malakoff, France.
- . n.d. "Describing Social Rhythms with Optimal Matching." Unpublished manuscript. Sciences Po, Observatoire Sociologique du Changement, Paris, and Institut National de la Statistique et des Etudes Economiques, Laboratoire de Sociologie Quantitative, Malakoff, France. http://laurent.lesnard.free.fr/IMG/pdf/Describing_social_rhythms_with_optimal_matching-2.pdf (accessed March 25, 2010).
- Levine, Joel H. 2000. "But What Have You Done for Us Lately? Commentary on Abbott and Tsay." *Sociological Methods and Research* 29 (1): 34-40.
- McAllister, William, Li Kuang, and Mary Clare Lennon. 2010. "Re-thinking Research on Typologizing Homelessness." Working Paper no. 10-01. Columbia University, Columbia Population Research Center, New York.
- Muthén, Bengt. 2003. "Statistical and Substantive Checking in Growth Mixture Modeling: Comment on Bauer and Curran (2003)." *Psychological Methods* 8 (3): 369-77.
- . 2004. "Latent Variable Analysis: Growth Mixture Modeling and Related Techniques for Longitudinal Data." 345-68 in *The Sage Handbook of Quantitative Methodology for the Social Sciences*, edited by David Kaplan. Thousand Oaks, CA: Sage.
- Nagin, Daniel S. 2004. "Response to 'Methodological Sensitivities to Latent Class Analysis of Long-Term Criminal Trajectories.'" *Journal of Quantitative Criminology* 20 (1): 27-35.
- . 2005. *Group-Based Modeling of Development*. Cambridge, MA: Harvard University Press.
- National Alliance to End Homelessness. 2007. "Chronic Homelessness." Fact Sheet, April 24. National Alliance to End Homelessness, Washington, DC.
- Rohwer, Götz, and Ulrich Pötter. 1999. *TDA User's Manual*. Bochum, Germany: Ruhr-Universität.
- Rossi, Peter H., Gene A. Fisher, and Georgiana Willis. 1986. "The Condition of the Homeless of Chicago: A Report Based on Surveys Conducted in 1985 and 1986." Report, September. University of Massachusetts, Social and Demographic Research Institute, Amherst.
- Sankoff, David, and Joseph B. Kruskal. 1983. *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*. Reading, MA: Addison-Wesley.
- Sosin, Michael R. 2003. "Explaining Adult Homelessness in the US by Stratification or Situation." *Journal of Community and Applied Social Psychology* 13 (2): 91-104.
- Stark, David, and Balazs Vedres. 2006. "Social Times of Network Spaces: Network Sequences and Foreign Investment in Hungary." *American Journal of Sociology* 111 (5): 1367-1411.
- Stinchcombe, Arthur L. 1968. *Constructing Social Theories*. New York: Harcourt, Brace & World.
- Stovel, Katherine, and Marc Bolan. 2004. "Residential Trajectories: Using Optimal Alignment to Reveal the Structure of Residential Mobility." *Sociological Methods and Research* 32 (4): 559-98.
- U.S. Interagency Council on Homelessness. 2003. *The 10-Year Planning Process to End Chronic Homelessness in Your Community: A Step-by-Step Guide*. Washington, DC: U.S. Government Printing Office.
- Weisburd, David, Shawn Bushway, Cynthia Lum, and Sue-Ming Yang. 2004. "Trajectories of Crime at Places: A Longitudinal Study of Street Segments in the City of Seattle." *Criminology* 42 (2): 283-321.
- Wishart, David. 1969. "An Algorithm for Hierarchical Classifications." *Biometrics* 25 (1): 165-70.
- . 2004. *ClustanGraphics Primer: A Guide to Cluster Analysis*. Edinburgh: Clustan Limited.
- . 2005. "Number of Clusters." 1442-46 in *Encyclopedia of Statistics in Behavioral Science*, edited by Brian S. Everitt and David C. Howell. Vol. 3. Chichester: Wiley.
- Wu, Lawrence L. 2000. "Some Comments on 'Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect.'" *Sociological Methods and Research* 29 (1): 41-64.

Notes

This study was supported by a grant from the National Institute of Mental Health (no. R34 MH073651-01) and by the Center for Homeless Prevention Studies at Columbia University. We appreciate their support. We are also grateful to the Department of Homeless Services of the City of New York (DHS) for providing the data to conduct this study. The analyses and interpretations expressed herein represent the opinions and conclusions of the authors and do not necessarily reflect the opinions of DHS or its staff. We thank George Gavrilis, Tali Schaefer, Glyn Vincent, DHS staff, and participants in the Grand Rounds of the Center for Homeless Prevention Studies at Columbia University for their thoughtful comments.

1. See Culhane et al. (2007) for the application of similar theory and techniques to studying homeless families. This article focuses on Kuhn and Culhane (1998) because other researchers use their approach or cite their research (see, e.g., Goering et al. 2002; Caton et al. 2005; Kertesz et al. 2005, among many others), and their typology is commonly referenced in policy-making discussions (e.g., U.S. Interagency Council on Homelessness 2003; National Alliance to End Homelessness 2007; Cunningham 2009, among many others). For a review of homeless typologies in general, see Jahiel and Babor 2007.

2. The homelessness research cited by Kuhn and Culhane (1998) theorized different kinds of homelessness over people's lives, not just shelter experience over a few years. This article follows Kuhn and Culhane in empirically analyzing typologies based on homeless shelter use and not on people's broader experience of homelessness. However, the time-patterned approach of this article and the specific method it uses are not limited to shelter-based data. They are generally useful for constructing time-based homeless typologies however homelessness is conceptualized and measured.

3. Because this article is not interested in how homelessness has or has not changed over time, the difference in time periods is irrelevant. The present study uses a later time period because its authors have a substantive interest in analyzing more current homelessness.

4. Among these are 1,864 right-censored cases that remained sheltered past the end of the 3-year observation period. This censoring is not relevant for the current analysis because, following Kuhn and Culhane (1998), it is only interested in groups that can be formed from data over the 3-year period.

5. Because this study uses shelter data, it does not capture people's street homelessness, the homelessness of living doubled up or tripled up, the homelessness of living in SROs (single-room occupancy locations), or other definitions of homelessness. Thus, any typologies it produces are, empirically, typologies only of New York City shelter use at the time period of the data. Of course, this is true by design: the empirical work of the study is to produce typologies that can be compared to the New York City shelter-based typology of the Kuhn and Culhane (1998) study and to the data and method used in that study.

6. These measures are likely much less reliable than the temporal measures, as they are based on shelter users' self-reports at the time of shelter entrance and are recorded by administrative personnel untrained in data collection.

7. Other methods are also used to identify time-patterned typologies. Group-based modeling (Nagin 2005) and general growth mixture modeling (Muthén 2004) were developed and are used to analyze developmental theories about criminal behavior (Nagin), educational outcomes (Muthén), and other phenomena. These approaches differ from OM primarily in that they are semiparametric (Nagin) or parametric (Muthén); OM is nonparametric. (See Abbott 2001 for a discussion of parametric and nonparametric approaches. Reporting other differences goes beyond the scope of this note.) This study uses OM mainly because the large number of cases and time points in the data set create the opportunity to identify acutely varied patterns. The authors' experience in using all these methods suggests that OM better identifies such variety. For critiques and discussion of group-based modeling, see Eggleston, Laub, and Sampson (2004) and response by Nagin (2004), among other such critiques and discussions; on general growth mixture models, see Bauer and Curran (2003) and response by Muthén (2003), among other responses therein.

8. These deviations are squared Euclidean distances. The Ward's algorithm in the clustering software used in this study is able to treat the Levenshtein distances produced by optimal matching as Euclidean distances (see Wishart 1969).

9. This solution is validated by drawing another simple random sample of 5,000 people from the full 40,169-person population and carrying out the same analysis. The results identify a time-patterned three-group solution that is similar to the one reported for the main analyses, though the authors know of no statistical tests for differences.

10. For instance, case A-1 was sheltered for 27 days the first month and for 10 days the second month but was not ever sheltered again. This history is consistent with expectations for transitional homelessness. Histories for cases A-2, A-3, and A-4 are also consistent with these expectations and are included in the group of illustrative cases from which the transitional sample is drawn.

11. Specifically, the study disproportionately samples the population of illustrative cases in each group, and the sample size is determined by the smallest number of illustrative cases available for any one group. This is the episodic group, for which the sample is 84 cases. The dissimilarity matrix is n by m , where n is all members of a group, and m is the sample of illustrative cases. That is, $m = 84$ for all groups, and $n = .745 \times 5,000$ for group A; $.196 \times 5,000$ for group B, and $.061 \times 5,000$ for group C. In addition, because the observed dissimilarities are between relevant pairs of cases, the averages are calculated using the number of pairs of relevant cases in each group. This makes the denominator $n(n - 1)/2$.

12. In their study, Kuhn and Culhane (1998) find that the transitional group is 81 percent of the sample, has 1.4 shelter episodes on average, and is sheltered on average for 57.8 days. The episodic group is 9.1 percent of the sample, has an average of 4.8 episodes, and is sheltered on average for 263.8 days. The chronic group is 9.8 percent of the sample, has 2.3 episodes on average, and is sheltered on average for 637.8 days.

13. Total sum of squares is the total within-group sum-of-squared deviations from the cluster centroid (cluster average) for the one-group solution (i.e., when all cases are grouped into one cluster). The measure in the table equals group sum of squares/group n . The same logic is used to calculate total dissimilarity. Note that this is not the same total sum-of-squares statistic that is used for the analysis reported in table 2, though they are related.

14. How the time-patterned approach captures frequency and duration creates greater opportunities for variability than the opportunities created by the time-aggregated approach. Thus, the two methods are not analyzing the same variability. Therefore, it is not correct to say that the time-aggregated approach is explaining more of the same variability as the time-patterned approach. As a result, the study does not compare homogeneity across the two approaches but only differences in homogeneity across solutions within approaches.

15. A 95-percent interval produces more statistically significant odds ratios in each of the two approaches.

16. Kuhn and Culhane (1998) do not explicitly theorize that age is a factor in the transitional-episodic relations, but the transitional group is found to be younger than the episodic group. Neither do they explicitly theorize that the transitional category is related to having any problem or to having all three problems. Such theory, however, seems to follow from the logic of their argument and is empirically explored by them. What these latter relations should look like for the chronic-episodic relationship is not clear, because chronically homeless people are expected to have more medical problems than people in the episodic group, and episodically homeless people are expected to have more mental health and substance abuse problems than people in the chronically homeless group.

17. To be clear: the extant research on types of temporal-based homelessness, not Kuhn and Culhane's (1998) expression of this research, was theoretically weak. McAllister, Kuang, and Lennon (2010) describe more extensively than the discussion does here some theoretical and empirical problems with the typology arising from that research.

18. An 18-group solution also meets these criteria but can be organized into the same four substantive categories as the 10-group solution and only slightly improves homogeneity (e.g., total explained sum of squares is 76 percent rather than 71 percent). The 10-group solution is reported because it is more parsimonious, although the 18-group solution may be more useful for some purposes (e.g., a services program may want greater within-category delineation).

19. These results support the thrust of Kuhn and Culhane's (1998) findings that there are three typological categories when a time-aggregated approach is used. The purpose of this further time-aggregated analysis is to see if solutions with more than three groups

produce results that are substantively different from the time-aggregated three-group solution and substantively different from the time-patterned 10-group solution. Thus, improvements in explained variance for time-aggregated solutions with larger number groups are not relevant and are not reported. Also, as noted previously, because the time-aggregated and time-patterned approaches create different total variation, comparing explained variance across these approaches is not useful. (These different total sums of squares for each approach are reported in table 4.)

20. This study discusses the causal importance of structural conditions because those conditions select certain people from a population of individuals. Although it is not impossible that individuals intentionally select circumstances from a population of structural conditions, it is less likely. People may, for instance, choose to move between towns to take advantage of local conditions (e.g., guaranteed shelter).

21. Suggestive explanations could also be made for the two other categories. This discussion does not try to explain why people initially enter shelters. The empirical analysis and, hence, the typology, is based on histories of shelter use; it does not incorporate people's lives before they first enter shelter.

22. Strictly speaking, shelter environment is an explanation relative to what a person imagines his or her nonshelter living environment would be. Use of the term "shelter environment" is shorthand. It implicitly suggests that, for those who remain sheltered, the environment is more valued than that outside shelter. Some individuals with substance abuse problems, for instance, may prefer the institutionalized housing of shelters to the institutionalized housing of jail.

23. Here and in the remainder of this section, the discussion of individual traits is not intended to focus on them but to draw attention to structural conditions and to the relationship between individual qualities and structural conditions. Because data on structural conditions are not available, the discussion attempts to glimpse them through the combination of the kinds of people they select for and people's shelter histories. The language of "selecting for" argues that conditions are causal; the language of "people choosing" argues that individuals are causal. The authors think that both causes are operating, but have no theory or data on how conditions and individuals interact to produce shelter histories.

24. The number of groups is a shorthand way to also discuss group size and nature. Choosing the number of groups also determines their size and nature, because choosing one number of groups over another means a different size for at least one group and introduces at least one group with a different character. Thus, discussions about number are not just about a cardinal number but about the character of the concepts described by the typology.