Analyzing Life Course Patterns with the Interval Graph Approach*

Joy E. Pixley**

Cornell University

Carter T. Butts***

Carnegie Mellon University

June 2001

Running Title: Interval Graph Approach

Word count: 11,758

* Please address correspondence to the first author at the Cornell Employment and Family Careers Institute, G19 MVR Hall, Cornell University, Ithaca, NY 14853 or electronic mail at jep23@cornell.edu. The Cornell Community Study was conducted by the Cornell Employment and Family Careers Institute (with acknowledgements to Phyllis Moen, Stephen Sweet, and the Cornell Computer Assisted Survey Team, especially Yasmin Diciccio, Lisa Horn, and Michael Miles) and sponsored by the Alfred P. Sloan Foundation (Sloan FDN #96-6-9, #99-6-23, Phyllis Moen, Principal Investigator).
** This work was supported in part by the Cornell Employment and Family Careers Institute.
*** This work was supported in part by the Center for Computational Analysis of Social and Organizational Systems and the Institute for Complex Engineered Systems at Carnegie Mellon University.

Analyzing Life Course Patterns with the Interval Graph Approach*

June 2001

Running Title: Interval Graph Approach

Word count: 11,758

ABSTRACT

This paper illustrates the *interval graph approach* to studying the life course. Under this framework, life histories are represented as networks of relations among spells (e.g., jobs, marriages, children, schooling). The focus is thus on the intersection of roles across time. We introduce two potentially important uses of this approach: to provide parsimonious summary measures of multi-domain life histories; and to enable comparisons of individuals' overall role overlap patterns over the life course. Applications of these methods are demonstrated using retrospective life history data from a sample of 1,532 men and women in Upstate New York. Implications of methodological variations are discussed and some potential uses of the methods for life course researchers are suggested.

INTRODUCTION

Scholars of the life course, including many who focus on work careers or family processes, face the challenge of assessing multiple role trajectories over time. The life course perspective motivates interest not only in the characteristics of individuals' current work roles, for instance, but also in events within their other, coterminous roles, and in the pathways that led them to their present situations (e.g., Elder 1995). This viewpoint has generated stimulating discussions on how ordering, spacing, and grouping of roles and events can affect individuals' choices and experiences, as well as on the sets of "typical" trajectories which emerge within particular socio-temporal contexts (Elder, Shanahan, and Clipp 1994; George 1993; Giele 1998; Moen, Elder, and Luscher 1995). A wealth of panel studies and studies incorporating life history reviews provides the requisite data for such analyses. This same data, however, has drawn attention to limitations of traditional methodological approaches in this area. Despite the richness of many life history data sets, a satisfactory implementation of the life-course approach has been hampered by the difficulty of analyzing complex, multidimensional, and internally related data with standard statistical tools (e.g., Singer et al. 1998). Attempting to grapple with life-course questions using the available methods can be like trying to use a two-dimensional representation to solve a three-dimensional puzzle. We are encouraged to look at experiences not only in one life domain, such as work or family, but across concurrent life domains as well. We are cautioned to consider not only current situations, but also the pathways across time that engendered those states. However, currently utilized methods are designed to focus either on relationships between roles at one point in time, or on the sequence of transitions in a single role over time. A technique for assessing the changing pattern of intersecting multiple life domains over time would be a useful addition to the current repertoire.

Fortunately, the field of life course studies is not alone in its need to represent and analyze complex, relational data. Social network analysis – at first blush an unlikely source for life course methods – has also had to deal with the challenge of analyzing relational data, and has developed a wide array of procedures for uncovering, describing, and predicting structural relationships. By identifying graphic (i.e., network) representations for life history data, researchers can apply established techniques for analyzing interpersonal relationships to the *intrapersonal* relationships between life course spells. Recent statistical innovations in network analysis provide methods for analyzing overall life course patterns that were unthinkable when many of the theoretical bases for life course research were first conceived, and the potential gain of applying these tools to life history data is substantial. The premise of the resulting new approach is that the life course can be meaningfully represented as a pattern of *relationships between roles or states*, based fundamentally on their overlaps over time (Butts and Pixley 2001). In this paper, we introduce two potentially important uses of this *interval graph approach* to life course research: to provide parsimonious summary measures of multi-domain life histories, and to enable comparisons between individuals' overall role overlap patterns over the life course.

Comparison with Existing Models

Before moving to a consideration of this new approach, it is important to review some of the other innovations that have been made in life course research over the past few decades. Arguably, the most promising methods currently available for studying life course patterns utilize event history analysis (e.g., Drobnic, Blossfeld, and Rohwer 1999; Long, Allison, and McGinnis 1993; Wu and Martinson 1993) or sequencing techniques (e.g., Abbott and Hrycak 1990; Blair-Loy 1999; Giele and Elder 1998; Han and Moen 1999). Event history models are primarily designed for studying predictors of specific transitions, such as changes in educational or fertility status. They are generally based on a shared stochastic modeling framework founded on an assumption of "rare events" (in the poisson sense) whose underlying rate of occurrence may be modeled by relating observed rates to various covariates. Sequence analysis, particularly in conjunction with optimal matching techniques, is more often appropriate for analyzing and comparing whole trajectories within a certain domain, such as work careers over time. The sequence analysts have tended to take a more data-analytic approach to studying the life course, and have placed less emphasis on model-based inference than have event history researchers. Though there is a lively debate about the relative utility of these methods for life course research (see, e.g., Abbott and Tsay 2000; Levine 2000; Wu 2000), we would argue that both approaches are appropriate, albeit for different research tasks. In the initial, exploratory phases of research (particularly when prior knowledge is weak regarding transition processes and associated covariates) purely data analytic, descriptive procedures, such as those most often associated with sequence analysis, are generally appropriate. Later, when additional knowledge permits specification of a plausible stochastic model, inferential approaches, such as those of event history analysis, become more suitable. The interval graph approach to life course research (which we employ here) entails both a representational scheme and an associated family of methods for analyzing and comparing life course patterns (Butts and Pixley 2001); in this respect, it is similar to both event history and sequence analysis. The specific methods introduced here are data analytic techniques which in some ways resemble and, we believe, complement work currently being done using sequence analysis. However, this general approach is not limited to exploratory procedures, and future work will elucidate the related inferential techniques.[1]

The Interval Graph Approach

One characteristic common to event history and sequence analysis methods is that their strengths lie in studying changes in a single domain over time. While it is possible in certain circumstances to adapt these methods to include simultaneous changes in other domains, such solutions tend to be inelegant at best. At worst, the consideration of multiple domains can render analyses computationally infeasible and require a considerable number of additional modeling assumptions. Even Andrew Abbott, a stalwart supporter of sequence analysis, recognizes that its unilinear nature is an "obvious drawback" for certain applications (Abbott and Tsay 2000: 10).[2]

In the interval graph approach to the life course, life histories are represented as a pattern of overlapping life course spells (e.g., jobs, marriages, children, schooling). The focus is thus not on the order of events (although this can be incorporated; see Butts and Pixley (2001)), but on the co-occurrence of roles or states. Elements of the life history are treated as relational entities, and the structure of connections between elements is the primary object of study. As such, this approach is well-suited for examining interactions between multiple domains, and can address research questions for which a sequence analytic or event history approach is not as appropriate. Although individuals' roles may be related to one another in various ways, the co-occurrence of roles is a key theme in many areas of research. The study of work and family roles, for instance, often indicates that the nature of working, being married, or having children differs depending upon whether one holds either or both of the other two roles at the same time (e.g., Han and Moen 1999; Hotchkiss and Moore 1999; Spain and Bianchi 1996). The study of life trajectories suggests that current roles may be better understood by examining the pathways taken up to that point, including, in the terminology of the interval graph approach, the pattern of overlapping roles held previously (Moen 1997; Shanahan 2000; Spain and Bianchi 1996).

The inclusion and definition of types of life course elements (most often spells) to consider in representing the life history is guided by the research question. Researchers who wish to include employment spells, for instance, may choose to define such spells as any periods in which individuals work for pay, or define part-time and full-time jobs as different types of spells, as prior knowledge dictates. Other researchers may choose to include spells based on physical disability, involvement in elder care, child custody, volunteer work, or geographic location. In all cases, the needs of the researcher and the availability of the data should motivate the choice of spell representation (see Butts and Pixley (2001) for further discussion).

Here, we discuss two families of analyses enabled by this approach to life course research: those involving comparisons of descriptive indices on life course graphs, and those involving direct comparisons of complete life history structures. In the first, patterns of overlaps between spells are assessed using interval graph techniques, which provide non-parametric indicators of individuals' life course characteristics. This procedure provides new, more parsimonious ways of describing data about life course patterns without making any assumptions about internal causality. In this paper, we focus on graph-level indices, such as the overall density of overlaps or the number of overlaps between employment and child spells.[3] These indices can then be used in any model suitable for descriptive measures.[4] For instance, they could be used as independent variables in a regression model (in conjunction with other indicators, if desired, such as characteristics of childhood background or timing of early transitions) to predict current outcomes of interest.

In the second family of analyses, the construction of quantitative differences between interval graphs enables the direct contrast of individuals' overall life paths. A structural distance analysis,[5] made possible by recent innovations in interstructural comparison methods (Butts and Carley 2001), calculates the distance between each pair of life course patterns. The distance score indicates the number (or in some specifications, the aggregate length) of each type of overlap, such as job-child or child-child, that would need to be changed in order to make two graphs identical. A variety of analyses can be conducted with structural distances. In one, multidimensional scaling of the inter-graph distances can be used to uncover latent dimensional structure, particularly in conjunction with procedures such as canonical correlation analysis or multivariate regression. Such analyses can be used to reveal, for instance, which life experiences correlate most strongly with key dimensions of overall distance between life course graphs. Alternately, analyses can focus on distances between specific pairs of people, such as spouses. This would be appropriate for questions such as the extent to which spouses with more dissimilar initial work experiences later exhibit more dissimilar life course patterns, or whether similarity of life course patterns is related to perceived spousal support. Similarly, an "analysis of distance" (analogous to an analysis of variance) can assess within- and between-group differences in life course patterns among, for example, professional women, professional men, nonprofessional women, and nonprofessional men. Finally, cluster analysis techniques can reveal groups of similar graphs, allowing a classification of typical life course patterns. The appropriateness and effectiveness of each of these applications rest heavily on the substantive theoretical assumptions, particularly as they affect how life course spells are defined and the range of variation on key indicators within the selected population.

INTERVAL GRAPH METHODOLGY

The interval graph approach to the representation and analysis of life histories is based upon graph-theoretic techniques that are already well established in the study of social networks. Standard methods exist for defining and calculating both characteristics of individuals within networks and network-level indices (Wasserman and Faust 1994). Typically, these techniques are used to describe in detail a single network, such as a large organization, or to compare networks using graph-level indices, such as size of network or density of ties. Seeking to compare tie patterns of different networks, Banks and Carley (1994) suggested contrasting the adjacency matrices of each network (a matrix that indicates whether each pair of vertices has an edge, or tie, between them) using Hamming distance (an established method of comparing sequences). Recent work building on this idea has rendered it practicable for partially labeled or unlabeled structures and extended it to the modeling of graph covariance, enabling the application of standard multivariate methods to the comparison of entire structures, such as social networks (Butts 1998; Butts and Carley 1998; Butts and Carley 2001). This technique is only now being refined to apply specifically to the study of life course patterns (Butts and Pixley 2001). While the technical aspects of this approach are being introduced concurrently to a audience familiar with similar methodologies, this paper is the first to present it to a broader range of social scientists and to include examples of data analyses.

Here, the basic elements of this technique are described and illustrated, to an audience assumed to have little or no prior experience with graph theory or network methods.

(FIGURE 1 ABOUT HERE)

Simple Graphs

A life history can be represented as spells situated on a time line, as in Figure 1. Much of this woman's life can be gleaned from this chart. After finishing school, she began a full-time job. Shortly after marrying, she changed jobs. She continued working at this job after her first child was born, but quit when her second child was born. After about a year, she took a part-time job, but when the marriage ended, she increased her work hours, and had three full-time jobs over the next fifteen years. A couple of years after both children had moved out, she began a new marital relationship.

Thus presented, the life history chart clearly helps to illustrate narrative accounts of people's lives. However, it offers no means with which to compare one narrative to another that are not reducible to spell-level characteristics, such as comparing the number of jobs or children, or the amount of time spent in full-time versus part-time employment. In the life course perspective, lives are more than simple aggregates of micro-level measures. The challenge, then, is how to compare life course patterns *in toto*, preserving key concepts revealed in the life history chart, such as ordering and co-occurrence of spells. To do so, we represent the life course pattern using graph-theoretic techniques.

(FIGURE 2 ABOUT HERE)

Newcomers to network analysis may benefit from an overview of the terminology, which is easiest to understand when illustrated by a simple graph (see example in Figure 2). In this graph, the vertices and the ties between them could represent linkages between a wide range of objects or concepts. As most readers will be familiar with the construct of a social network, let us imagine this graph to be a sociogram. Each vertex thus represents an actor, while each tie is a relationship between those two actors. One notable aspect of the graph in Figure 2 is that it is not "connected." That is, there are three *components*: while most actors are connected in one component, v_a (vertex "A") is unconnected to any other actor, and v_k and v_j are connected only to each other.

As network characteristics are generally only meaningful for groups of three or more vertices, we confine our illustrations of these indices to the larger component. Centrality is a vertex-level property which can be measured in a variety of ways, three of which we offer here. In this group, v_d, v_f and v_g are each connected by ties to five other actors: they exhibit the highest *degree centrality*, while v_b and v_l exhibit the lowest, each being tied to only one other actor. The *betweenness centrality* index indicates how frequently the vertex lies on the geodesic, or shortest path, between two other vertices.[6] Here, for example, v_b, v_c, v_k and v_l have a betweenness score of zero: put simply, there is no pair of actors for whom any of these actors function as a bridge on the shortest path linking the two.[7] Note that v_d and v_g have higher betweenness scores than v_f, even though they are connected to the same number of people, because the first two each know someone who is not connected to anyone else in the group. So, whereas most of the people v_f knows also know each other, the shortest path between v_b and another actor is always through v_d (and likewise for v_g and v_l). A third type of centrality is *closeness*, which is calculated using the average length of the shortest path (the number of intervening actors) between an actor and all other actors. (While this is ill-defined for disconnected graphs, recall that we are here considering only the largest component.) Again, v_d, v_f and v_g are the most central in this respect, being directly tied to five other actors and only one actor away from the remaining two, while v_b and v_i have the lowest closeness centrality.

Graph-level indicators of *centralization* indicate the extent to which degree, betweenness, or closeness centrality is concentrated on one actor. The highest score is obtained when one actor has the maximum possible centality and the rest of the actors have the lowest possible centrality,[8] while the lowest score is obtained when all actors have equal centrality. The *density* of a graph is the number of social ties as a proportion of the total number of possible ties, given the number of actors. While the graph in Figure 2 may look dense, it is easy to imagine an increase in the density if more pairs of actors were linked to each other (e.g., connecting v_a, v_j and v_k to other actors in the main component). As it is, 13 ties are realized out of the possible 55 dyads, for a density of .24. Other graph-level indices of interest include the *median degree* (the median degree centrality score), the *diameter* (the longest geodesic, or shortest path between two actors, in the graph), and the *mean geodesic distance*.[9]

Simple Interval Graphs

The same indices just described apply also to interval graphs, although the meaning of the indices can be less intuitively obvious. An interval graph can be constructed with vertices that are not discrete points (e.g., actors) but intervals (e.g., the interval of time a computer in a network will be busy). Here, ties refer to overlaps between those intervals (e.g., when two computers will be busy at the same time). Interval graphs are thus useful for analyzing coordination problems within a single system, such as scheduling networked computers, employees, or circuits.[10] They can also be used to examine patterns of life course spells, which are treated as time intervals spent in certain roles.

(FIGURE 3 ABOUT HERE)

An example of a simple life history interval graph is shown in Figure 3, in which the pattern of ties shown in Figure 2 is seen to represent the overlaps between spells in the life history chart of Figure 1. As the figure indicates, overlapping intervals are roles held at the same time. The three components of the graph are now referred to as different *life epochs*. This woman's educational spell, for instance, is unconnected to any other roles she holds later in life: in the few years after she graduates, literally every domain in her life changes. Then again, almost thirty years later, she is working with colleagues and living with a partner who did not know her when she was married, or working part-time, or when her children were still living at home. In between those two epochs is a large, connected life epoch in which some part of her life always stays the same, even when other parts change. A lack of connection between spells does not necessarily indicate that earlier spells have no influence on later spells: e.g., an educational degree should still influence later work experiences, even when those spells do not overlap. However, connectedness in the form of life epochs is likely implicated in the continuity (and discontinuity) of many aspects of people's lives, such as geographic mobility, stability of social networks, and identity transitions.[11] Similarly, life epochs may indicate periods where multiple roles buffer the stress of role transitions in other domains.

In a life course graph, degree centrality indicates the number of ties to other spells. In this example, the three spells that have the highest centrality are also, not surprisingly, the longest spells: the first marital spell and the two child spells. The partner and each child overlap with each other and with three job spells each, forming an epoch which links all but the last of the woman's jobs. The part-time job has the highest degree centrality of the jobs, as it is held while all three alters are living with the subject, whereas she is living only with her first husband during her first full-time job, and only with her youngest child during her fourth full-time job. These two jobs, although they are in the same epoch, have the lowest closeness: they are linked to other spells in this woman's life more remotely than the middle three jobs. At the graph or component level, the diameter and the mean geodesic distance similarly indicate the extent to which spells are linked together.

The graph indices described so far were developed primarily with social networks in mind. Presumably, applying graph theoretic techniques to a new problem, such as life course analysis, should encourage development of revised or new indices specifically appropriate to the topic. As a case in point, while the actors in the sociogram were treated as equivalent vertices, the spells here represent different types of experiences. As such, the substantive meaning of an overlap depends upon the spells involved. Furthermore, overlaps between certain spells are expected to be common (e.g., overlaps between child co-residence spells) while others may be impossible by definition (e.g., overlaps between legal marriages). In this vein, we introduce two additional types of density, which are based on block modeling techniques.[12] Recall that density typically refers to the number of ties in a graph as a proportion of the total number of ties possible, given the total number of vertices. Defining density more narrowly within a "block" limits which kinds of ties are included in this calculation. By specifying certain spells as belonging to the same block, or domain, *intra-domain density* is calculated as the proportion of overlaps between spells of the same domain (e.g., overlapping part-time jobs) to the number of possible overlaps (given the total number of part-time jobs). Similarly, *inter-domain density* is the proportion of overlaps between spells of different domains to the number of possible overlaps. Thus, if two graphs have the same number of child spells and job spells, but the second graph has a higher child-job density, this indicates that the second has more overlaps between these two types of spells. Both of these density measures are valid only if enough of the relevant spells exist to form at least one overlap (e.g., two job spells, or at least one job and one child spell).

As with any density measure, intra- and inter-domain density measures assess both the number of the type(s) of spell(s) involved and the number of overlaps between those spells. For example, one graph would have a lower intra-domain density for child spells than another if it has the same number of child spells but fewer overlaps between them, or alternately, if it has the same number of overlaps but more child spells. As such, density should be assessed in tandem with other indices (such as number of spells of each type) to facilitate interpretation. Intra-domain density is expected to differ by domain types. As siblings tend to live with each other, children spells often overlap, while intra-domain overlaps should be less common among full-time jobs, educational spells, and marital spells. Low intra-child density or high intra-full-time job density thus indicate unusual life graphs.

(FIGURE 4 ABOUT HERE)

To illustrate the density measures, two graphs are shown in Figure 4 that have the same number of spells with somewhat different patterns of overlaps. Here, the differences are due to the continued co-residence of three children after the dissolution of the first marriage in Graph A, whereas the three child co-residence spells in Graph B are split between two marriages. Graph A has an inter-child density of 1.00 and a marital-child density of 1.00: all possible overlaps between child spells and between child and marital spells are observed. In Graph B, only the first two child spells overlap, for a inter-child density of .33, and there are only three marital-child overlaps (of six possible) for a marital-child density of .50. Similar differences are seen in the scores for inter-domain density between child and full-time job spells.

The relationship between certain types of inter-domain and intra-domain density can be constrained by the substantive nature of the spell types. For instance, note that marital-full-time job density is much lower in both graphs (.30) than it would be in a graph that had only one marital spell which extended from the beginning of the first spell to the end of the time line (.60). This is to be expected for these types of spells. As both marital and full-time job spells tend to exhibit low intra-domain density, a second marital spell is likely to overlap with, at most, one of the jobs that overlap with the first marital spell, reducing inter-domain density considerably. If the first marital spell in Graph A were stable across the life course, the same inter-domain density of .30 could only be obtained if there were eight additional jobs prior to the marital spell. Another notable point is that, as child co-residence spells are likely to overlap, stability of marriages that overlap with the child spells will have little effect on child-marital density for the parent who retains custody of the children (as in Graph A) but can strongly affect it for the parent who does not (as in Graph B).

Once researchers have constructed graph-based indices of life course patterns, such as those mentioned here, they can be used to describe and compare individuals using a wide range of standard analytic techniques. Selected illustrations of such applications are included later in this paper.

Interstructural Analyses

A second family of analyses involves directly comparing graphs based on the pattern of ties they exhibit. At the most basic level, the *distance* between graphs is the extent to which each graph would have to be altered in order for the two graphs to exhibit the same pattern. The particular specification of the distance calculation, however, is by no means singularly determined, and must be defined by the researchers based upon their substantive interests and assumptions. At this early stage, only preliminary evidence exists concerning the relative merits of specifications. Furthermore, it is likely that the appropriateness of specifications varies across research questions.

Three key specifications are briefly discussed here: what it means for two graphs to exhibit the "same" pattern (exchangeability, or labeling, of vertices); what values should be assigned to particular ties between spells; and the measure used to assess "extent of change" (choice of metric). We consider each of these in turn. In the special case of uniquely labeled graphs, vertices are unique and cannot be exchanged for other vertices. This requires exact matches between sets of paired vertices for structures to be considered "the same": for instance, if v_a and v_c exhibit a tie in the first graph, v_a and v_c must also exhibit a tie in the second graph to constitute a match. Using labeled graphs is useful when comparing different types of ties within the same vertex set: for example, assessing whether advice ties map onto friendship ties between actors in a single social network. However, this is often inappropriate when comparing graphs defined on different sets of vertices, such as multiple social networks, where vertices on one graph cannot be labeled such that they are equivalent to specific vertices on the other. Unlabeled graphs – the opposite extreme – make no distinction between vertices, allowing total exchangeability. Actors, while retaining their ties to other actors, may be rearranged so that the *pattern* of ties most closely resembles that of the other graph, with "the same" pattern then reflecting structural isomorphism.[13] Between these two extremes, in partially labeled graphs, vertices are considered exchangeable within but not between defined types: for example, managers may be exchanged with other managers, but not with secretaries. Conceptually, partial labeling appears to be the most appropriate for comparing life course graphs. This specification allows permutations of tie groupings within domains, such as switching the order of two job spells so that their overlaps with other spells more closely match the pattern of another graph, but does not allow exchanges across domains.

Given a notion of what pattern allows two graphs to be considered "the same," we now move to a consideration of how ties between life elements are to be valued. The most basic value that ties between vertices can take is a dichotomous indicator of their existence (presence or absence). Here, the *simple life history graph* specification reflects only whether pairs of spells overlap. Comparing life history graphs on this relatively simple dimension may be preferable to more detailed graphs. Given the great variety of life course pathways, elementary expressions of those pathways may be more likely to elicit indicators of commonality. However, additional complexity can be incorporated by associating a value with each tie which describes some characteristic of the relationship between intervals.

In prior work, we developed four types of tie values for use with life history interval graphs (Butts and Pixley 2001). The *raw spell overlap* is defined as the length of time the two spells overlap. Unless otherwise noted, the term "overlap graph" refers to raw spell overlap calculations. The *directional overlap* is defined as the length of time the two spells overlap if they overlap, or the length of time separating the two spells (expressed as a negative value) if they do not overlap. For the *fractional overlap*, each pair has one value defined as the proportion of the length of v_i that overlaps with v_j and another defined as the proportion of the length of v_j that overlaps with v_i. The *fractional joint overlap* is defined as twice the overlap divided by the sum of the length of v_i and the length of v_j.

Finally, the choice of metric determines how distances are calculated. Fifty years ago, Hamming (1950) introduced a simple metric to compare sequences of communications: the total number of places on one sequence where the paired sequence did not exhibit a match constitutes the distance between sequences. While calculating distances between graphs is quite different from comparing sequences, the term *Hamming distance* is still used to refer to this count metric. In interstructural comparisons, Hamming distance is the total number of ties that would have to be changed in the graphs to produce matching graphs (under the definition of "matching" considered above). Other metrics assess not just whether points or ties on two structures differ, but the extent of the difference. As such, they are appropriate for comparing graphs specified with valued ties, such as raw overlap graphs. The metric of absolute difference arguably requires the fewest assumptions, simply summing the absolute differences between matched overlaps to determine the distance between graphs.[14]

Distances between simple life history graphs are expressed as the number of overlaps that would have to be changed to make the graphs equivalent (note that this is true whether Hamming or absolute distance is used because the differences are expressed as counts rather than amounts[15]). Distances for raw overlap, directional overlap, fractional overlap, and fractional joint overlap graphs are expressed as the number of units of time (e.g., days, months) that would have to be changed to make the graphs equivalent. Note that tie values do not effect the graph-level and vertex-level indices described above, such as density and centrality.[16]

Evaluating the Effects of Tie Values on Interstructural Distance Measures

If an independent set of criteria for determining differences between life course patterns existed, we could use it to assess the relative accuracy of representing graphs using different tie value specifications. The representation that most closely reproduced the differences described by such criteria would be judged as most appropriate for use in life course research. However, as a field, we have yet to define what are considered to be meaningful distinctions between life course patterns with enough precision to produce such criteria. Furthermore, life course researchers may not agree on such judgements, particularly in the initial rounds of such an attempt. It is simple to look at three life history graphs and describe the ways in which they differ. However, it is relatively difficult to determine whether the overall difference between person A and person B is, say, at least twice that of the difference between person A and person C, as this requires establishing which differences to weight more heavily than others. Although descriptive, case-by-case assessment is often performed in small, in-depth studies, an objective and consistent criteria is required if we are to compare large numbers of life course patterns.

As an initial step toward assessing whether the distance measures produced here match prior research and our intuitive understandings of difference, twelve life course graphs were created that represent substantively different groupings and orderings of spells. Descriptions of these prototypes are offered in Table 1. We calculated interstructural distances using simple life history graphs as well as graphs with valued ties, including overlaps, directional overlaps, fractional overlaps, and fractional joint overlaps. Cluster analysis using these distances produced theoretically coherent groups:

clusters based on simple life history graph distances are shown in Table 1. Most tie values exhibited similar results, although directional overlap graphs showed a somewhat less intuitive clustering of cases. Pair distances calculated using simple life history graphs correlated with those calculated for fractional and joint fractional overlaps at .98 and .96, respectively, overlaps at .83, and directional overlaps at .78. A principal component analysis indicates that ninety percent of the variance in pair distances is accounted for by the first component, onto which all five measures load almost equally (.42 to .47). This suggests that, with the possible exception of the directional overlap graph, the more complicated distance measures using valued ties do not produce substantially more information than using the simple life history graph. Although more thorough testing is warranted, the directional overlap graph appears to be prioritizing different, non-intuitive differences and similarities between graphs. Based on these analyses, distances using the simple life history graph are judged, at least preliminarily, to be most appropriate. However, we expect inferences based on distance structures to be reasonably robust to choice of graph representation.

(TABLE 1 ABOUT HERE)

Using a method that focuses on overlaps between spells, of course, already assumes that certain differences between life course patterns are more relevant than others for the associated analysis. Clusters of the sample "lives" illustrate these assumptions (see Table 1). The primary differences between clusters center around the types of spells each group has (or does not have) and their overlaps; timing, length, and number of spells have less impact. Some graphs were designed with family transitions in their early twenties, while others have relatively delayed entry into marriage and parenting following longer educational spells and "isolate" jobs. This type of timing effect is somewhat muted when it produces periods of isolates: jobs that overlap with no other spells look the same as educational spells with no overlaps. However, graphs with periods of isolate spells differ from other graphs to the extent that the other graphs exhibit overlaps that they do not.

Here, the three prototypical graphs with part-time jobs overlapping with child spells are grouped together (Cluster 1), as are the three graphs with no child spells, including one with a part-time job (Cluster 2), even though number of jobs and whether the first marriage ended varies. Two of the three graphs which include a second marital spell are grouped together (Cluster 3), while for the third, the presence of overlaps between part-time jobs and child spells (during the first marital spell) appears to pull it into the first group. The final cluster contains four graphs including children during a single marital spell and full-time job spells across most of the life course, although timing of family transitions differs, the marriage ends in two graphs, and one graph includes "isolate" (non-overlapping) extended education spells. On the whole, these distinctions make intuitive sense, suggesting the utility of distances calculated using the simple life history graph. However, before any strong claims could be made favoring one specification over the others, further testing is warranted, preferably utilizing exogenous criteria for expected differences between certain graphs.

ILLUSTRATION OF THE INTERVAL GRAPH APPROACH

Sample

We use data from 1,532 upstate New York residents who completed telephone interviews in the Cornell Community Study (CCS) to illustrate the analyses associated with the interval graph approach. The CCS used a household listing sample to locate households in selected New York Census tracts.[17] Married and cohabiting couples were eligible if at least one member was working at the time of the interview, and single individuals under the age of thirty were eligible if they were working or on family leave. The sample includes 1,260 couples in which both partners were interviewed, 255 partnered individuals whose partners were not interviewed, and 17 single individuals. The sample is divided evenly by gender. Most partnered individuals (95%) are married. Ages range from 21 to 79 years, with an average age for men of 46 years and for women of 44 years. The majority (90%) have had or lived with at least one child (including biological and non-biological children); parents average between two and three children ($\bar{x}$ = 2.6; s.d., 1.2). Most men (87%) are working full-time (35 hours per week or more), while a small proportion are working part-time (4%) or not at all (9%). Just over half of women (53%) are working full-time, with the other half evenly divided between those working part-time (24%) and those not working for pay (23%). Consistent with the middle-class neighborhoods sampled, incomes are high relative to national averages: the median annual income for working men is $51,500 ($\bar{x}$ = 64,238; s.d., 53,665) and for working women is $28,000 ($\bar{x}$ = 30,939; s.d., 21,302). The sample is relatively highly educated: the highest degree reported is a bachelor's degree for one-third (35%) and an advanced degree (primarily master's degrees) for one-fifth (20%) of the sample (no gender difference).

Retrospective life history information was gathered in a series of structured questions for the following domains: employment, education, marital/cohabiting relationship, children, and household members. This analysis focuses on the first four domains. Employment spells are separated into those involving full-time employment (35 hours per week or more) or part-time employment. Under education, all reported spells are used, whether or not a degree was received, including short educational spells such as certification programs. As the majority of relationships reported involve legal marriage, the term "marital" will be used to refer to all relationship spells. Spells involving children refer to co-residence with biological children and any non-biological children who lived with the respondent for at least six months. Children in this sample have between one and four separate spells of living with the respondent, including relatively rare instances of adult children returning home.

As indicated above, we will here demonstrate two families of analyses facilitated by the interval graph approach. First, parsimonious summary indices of graph-level characteristics are produced. These indices can be used with standard descriptive or comparative analytical methods, such as regression. Second, interstructural distances between each pair of graphs are calculated. The resulting distance matrix can be used in numerous ways, including MDS (multidimensional scaling), analyses of distance between and within groups or dyads, and division of subjects into clusters. We illustrate examples of each family of analysis below, beginning with descriptive analyses.

Graph-Level Characteristics

In this section, we present graph-level indices in relation to two important determinants of life course experiences: age and gender. First, for any group that differs in age, life course patterns should be affected by the exposure effect of the passage of time. Older individuals have had more time with which to accumulate life course spells and, presumably, to have overlaps between those spells. The correlation coefficients for the relationships between age and the graph-level indices produced using the CCS data are shown in Table 2. Age is seen to be positively related to number of spells and number of overlaps, as expected. However, broken down into type of spell, the relationship is only substantive for child spells (r =.39) and, to a lesser extent, for full-time jobs (r = .24). Although age is not related to number of epochs, it is positively (if weakly, at r = .20) related to median number of spells within epoch and diameter. Those in later midlife and beyond have had more time to amass larger groups of connected spells, leading to greater average distances between earlier and later spells. Median degree is also positively correlated with age, indicating that older respondents' spells generally have more overlaps to other spells than do those of younger respondents (although note that age is not related to density).

(TABLE 2 ABOUT HERE)

Age is also negatively related to the density of overlaps between child co-residence spells, indicating that older respondents are likely to have adult children move back home after the other children have left, or to live with additional children after their older children moved out. The latter suggests either families with large numbers of children who are relatively widely spaced, or "second" families begun with (or already started by) a later spouse. (Note that age also has a small but significant correlation with number of spouses.)

Similarly, age is negatively correlated to density of overlaps between part-time jobs and education, although age does not substantially increase the predicted number of either type of spell. As educational spells tend to be concentrated in early adulthood, younger respondents who worked part-time during college would have a relatively high inter-spell density, whereas older respondents who may have had part-time jobs in mid-life or after retirement should have relatively fewer overlaps.

(TABLE 3 ABOUT HERE)

A plausible criterion for any set of measures designed to describe life course patterns is that they capture known gender differences. Graph-level indices for men and women in this sample are shown in Table 3. Men average slightly more spells than women (11 versus 10) and more overlaps between them. Graphs for both men and women average between two and three life epochs, with a median of five spells each, suggesting that it may be common for graphs to have two epochs that contain about half the total spells each. The average density index is .37. Note that the example life history graph used earlier (Figures 1-3) has a density of .25; another seven overlaps would be needed to raise the density to the average indicated in this sample. The median degree, or number of other spells each spell overlaps, is somewhat higher, on average, for men than for women (3.30 versus 3.07). The mean geodesic distance and the diameter (i.e., longest geodesic) are both longer for men than for women.[18] Women exhibit slightly higher betweenness and closeness centralization, but degree centralization does not differ significantly by gender: this may be due to the sensitivity of the first two measures of centralization to the number of epochs (or the lack of connections between parts of the graph).

The overall picture can be summarized thusly: men have somewhat more spells than women within the same time period, which naturally results in more overlaps. At the same time, men's spells tend to have more overlaps per spell (median degree) and spells are further away from each other (e.g., diameter), while their graphs exhibit the same overall density. This suggests that men, like women, may generally have two or three primary life epochs, but have a higher density of overlaps within each epoch, which is attenuated by a third or even fourth epoch that contains only one or two spells. Women's epochs, by contrast, seem to cover more temporal territory with fewer links.

This would be consistent with one prototypical pattern in which women retain custody of children in the event of marital dissolution, and the longer children co-residence spells provide a link between otherwise unconnected jobs, educational spells, and relationships, whereas men are more likely to have "isolate" spells of full-time jobs that do not overlap with other roles.

These standard graph indices are consistent with gender differences, but more detail can be gained by examining specific types of overlaps between and within types of spells. Considering that an individual's children typically live together, density within child spells is expected to be quite high, which it is for both men and women. As men do not exhibit significantly more child co-residence spells than women, their slightly lower inter-domain density may point to women's greater tendency to retain custody of children at marital dissolution. When women have children with two different fathers (including step-children), those children would tend to live together, whereas when men have children with two different mothers, those children are unlikely to live together.

Men average more full-time job spells than women, while women average more part-time job spells than men. Men have overlaps between full-time jobs more often, while women have overlaps between part-time jobs more often, but they do not differ significantly in the density of these overlaps, given the number of relevant spells.

Differences in the number of part-time jobs and full-time jobs reflect women's greater propensity to work few or no hours when they are married, especially when they have children. The inter-domain overlap and overlap density indices, measures unique to the interval graph approach, clarify the story further. Men exhibit more overlaps between full-time jobs and both educational and family spells than do women, while women exhibit more overlaps between part-time jobs and both educational and family spells.

Furthermore, men have a higher overlap density between part-time jobs and both full-time jobs and educational spells than do women, while women have a higher inter-domain overlap density between part-time jobs and family spells than do men. So, even among men and women who have the same number of part-time jobs and the same number of family spells, more of the women's part-time jobs would overlap with family spells, while more of the men's part-time jobs would overlap only with full-time jobs or educational spells. The same is true for full-time jobs: even if men and women have the same number of full-time jobs and child spells, the proportion that overlap appears to differ substantially between the two groups. That is, men and women differ not only in how likely they are to have worked at full-time or part-time jobs at some point in their lives, but also in which other roles they held *while* they had those schedules.

On the whole, the number of spells and the types of overlaps are consistent with the differences in life course patterns by gender that we would expect to find. Many life course hypotheses could be tested using these graph-level indices, particularly the number and density of spell overlaps, either alone or combined with other characteristics of life course events, such as age at first marriage or first birth, highest educational attainment, or prestige of first job.

A simple demonstration

Once constructed, graph-level indices can be used as any other descriptive variables. Here, we illustrate this using regression. Given women's well-documented propensity to reduce their work hours when they have children, past overlaps between child spells and job spells could help to explain current income differences. Briefly, women who had full-time jobs when they had children may be selected into this position due to higher wage or career orientation, either of which would imply higher earnings later. On the other hand, we would not expect the same indicators to distinguish among fathers, as they tend to work full-time when they have children regardless of wage or attitude. For either gender, overlaps between full-time jobs and child co-residence could indicate continuity of employment, which should lead to greater average advancement. Overlaps between part-time jobs and child spells are not expected to have a clear relationship with later income, as the comparison group of those who did not work part-time during their child spell(s) includes both those who did not work at all and those who worked full-time. These predictions are tested using ordinary least squares regression.

(TABLE 4 ABOUT HERE)

Preliminary analyses indicated that a regression model pooling men and women, including the 389 couples in which both spouses have complete data, results in residu al errors with a small but significant correlation between spouses ($r = .13$, $p = .01$), indicating mild autocorrelation. Thus, models were run for men and women separately. Table 4 shows the unstandardized coefficients for the OLS regression of income (logged) on the standard predictors of work hours, education, and age.[19] These variables show significant coefficients in the expected direction for both men and women. In addition, graph-level characteristics were tested for inclusion into the best-fit model: those that added to model fit for either gender are shown here. After accounting for current education level, number of educational spells is negatively predictive of income for women. This is consistent with the notion that interruptions in the educational career are associated with lower rates of returns on subsequent educational attainment. Number of part-time job spells held in the past is negatively predictive of income, even after accounting for current work hours, but only among men. Number of child co-residence spells are negatively related to income for women, but not for men. As predicted, number of overlaps between full-time jobs and child co-residence spells has a positive relationship for women, but again, not for men. Although the effect sizes of the overlap indicator in this example are not large, this is not too surprising considering the breadth of other influences, particularly concurrent factors, on income. Such results are hopefully suggestive of similar research questions that could be addressed using graph-level indices.

Interstructural Comparisons

Next, we demonstrate a substantially different type of analysis that can be performed using the interval graph approach. Interstructural comparison involves calculating the distance between interval graphs, based on similarity of overlaps, as a measure of how different life course patterns are from each other. This produces an n x n matrix of pair differences that can be used in a variety of ways. For instance, differences within particular pairs, such as husbands and wives or siblings, can be examined. Groups, such as working women and housewives, can be compared to determine whether between-group differences in life course patterns are greater than within-group differences. Multivariate analyses may be used to uncover correlates of distance among other measures of interest. Finally, graphs can be clustered based on distance to indicate groups of people whose life course patterns are similar. We use the last method to illustrate the application of interstructural distances.

Cluster Analysis

Clusters based on interstructural distance indicate groups of people whose life course patterns are relatively similar to those of each other and dissimilar to those in other clusters. Initial analyses performed on the entire data set indicated that age was a key factor in differentiating clusters. However, this is properly seen as confounding, rather than elucidating, life course patterns. Individuals can be on the same life pathway but fall into different clusters if some have followed that pattern for forty years while others are just beginning to establish that pattern. Such a classification thus undermines the goal of identifying people who exhibit different patterns during the same periods in their lives.

The analyses presented here instead use patterns exhibited within each decade of life. Each life graph is divided into patterns occurring during the decades for which valid data is available, including the twenties, thirties, forties, and fifties.[20] Then, distances between decade sub-graphs are calculated and clusters determined within each decade. This allows for the fullest use of the data, while simultaneously distinguishing age effects (e.g., what happened in the subjects' twenties) from cohort effects (e.g., whether subjects were in their twenties in the 1950s or the 1990s). For instance, substantially later ages of marriage and child bearing, particularly when associated with additional educational spells, would suggest patterns of role overlaps in the twenties and possibly in the thirties that would differ from the early family formation pattern more common in the 1950s and 1960s. If a large cohort effect exists, some clusters would be comprised primarily of people from older cohorts while others would overwhelmingly include people from younger cohorts. A gradual cohort change is more likely, such that clusters contain some people from every cohort, although cohorts may be more or less prevalent in certain clusters. On the other hand, there is a strong potential for gender differences. Given women's relative propensity to not work outside the home or to work part-time, particularly when their children are young, it is expected that some clusters will consist primarily of women while others may be predominantly male.

Another life course question that decade-based clusters can elucidate is the possibility of pathways through clusters. Do people in one cluster distribute evenly across clusters in the next decade, or do they disproportionately end up together in the same cluster again? For instance, subjects who began their families in their late twenties would likely be clustered differently in the twenties decade than those with earlier parenting transitions; the obvious next question is whether they are clustered similarly in their thirties, since both groups will have child spells that span that decade.

As the purpose of this paper is broad in scope, the presentation of analytic results is necessarily constrained. Our primarily goals in this section are to demonstrate that inter-structured distances between life graphs can be used to classify life course patterns into meaningful clusters and to suggest some applications of this technique. Cluster analysis of graphs can be performed using distances based on any set of the specifications discussed earlier. Here, simple life history graphs are contrasted under partial labeling using the Hamming metric.[21]

Five to seven clusters were determined for each of the decades. There are only moderate indicators of cohort effects, but there are some relatively strong gender differences, which we discuss below. However, it is clear that, in terms of life course patterns, commonalties span both gender and cohort boundaries.

(TABLE 5 ABOUT HERE)

Descriptive characteristics by decade cluster are shown in Table 5. It is important to remember that these are individuals' *current* characteristics, not necessarily those they possessed over the course of that decade. To some extent, we can assumethat individuals shared many of the same characteristics at the time, such as marital status, number of children, and work status, or they would not be clustered together. It would be extremely difficult to present full information on individuals' characteristics during each period, as roles and the overlaps between those roles often change over the course of a decade. The age at first marriage and age at first birth (or adoption) of a child are included to give some indication of status at that time. While using current characteristics may complicate interpretation in some senses, it also facilitates what is arguably a more challenging test of hypothesized continuity in life course pathways: do people who are similar *now* tend to have taken the same pathway in the *past* – sometimes, decades earlier – to achieve the same situation?

Although a thorough description of these clusters is beyond the scope of this paper, some brief comments will hopefully illuminate some of the key differences. Some clusters contain more members from older cohorts: specifically, 20D, 30G and 40D. Members of these clusters have relatively low education, early marriages, early child-bearing, and large families (on average, four to five children). They are more likely to not be working than typical, possibly due to retirement, and average incomes are moderate. Those who are still working have moderate incomes for this sample.

One cluster in every decade is disproportionately female, while at least one is disproportionately male, and these generally map onto the classic breadwinner and housewife roles. Clusters 20B, 30F, 40B, and 50D are each comprised primarily of women. In general, individuals in these clusters married and had children fairly early, currently have two or more children, are relatively likely to be working part-time or not at all, and report the lowest average earnings (among those who are currently working).

Members of these clusters also tend to have lower educational attainment than average for this sample, particularly in the twenties and fifties.

Clusters 20A, 30A and 30E, 40A and 40C, and 50A and 50E are predominantly male. Individuals in Cluster 20A have a somewhat greater propensity to work full-time and about half have three or more children, but little el se distinguishes them from average. In all four clusters in the thirties and forties, these individuals are highly likely to work full-time. Those in Clusters 30A and 40A report the highest earnings of any cluster in those decades and are fairly well-educated, while those in Clusters 30E and 40C are more moderate on these dimensions. Clusters 50A and 50E do not differ as substantially from other clusters on income, probably due to the greater number who are not working or working part-time, presumably after retirement from their "career" job. Clusters 30A, 40A, and 50E contain people with large families of three children or more, whereas the other male-dominated clusters in those decades tend to exhibit two-child families.

Cluster 20C suggests the pattern of delaying family while building one's career. Relative to other clusters, it contains men and women with high educational attainment, high salaries, later age at first marriage and first birth, and a propensity to have few or no children. This cluster does not, as we might expect, include inordinate proportions of people from more recent cohorts. However, members of this cluster who are currently in their twenties and thirties are disproportionately likely to have only one or no children compared to older members. While it is possible that many of the younger members will add to their families in the future, a lasting cohort effect cannot be ruled out. Clusters 30C and 40F resemble 20C in terms of family formation, although the income levels are not as high. Cluster 40F also differs in terms of having fairly low educational attainment and a somewhat greater propensity than usual to have more than one marriage. Cluster 50C includes all individuals of that decade with no children, but is otherwise moderate, bearing little clear resemblance to the earlier low-children clusters. It is plausible that the absence of children drives the clustering so strongly, perhaps particularly in the older cohort, that the algorithm groups people who vary greatly in other characteristics.

This pattern of similarities lends support to the argument that clustering on interstructural differences between life course graphs produces meaningful groups. Given more space, a more thorough examination of the composition of these clusters would be preferred. In particular, much of the tripartite relationship between decade clusters, cohort, and current (that is, eventual) characteristics remains unexplored.

The next test assesses whether cluster membership forms patterns across decades. If knowing individuals' cluster memberships in one decade reveals nothing about their cluster memberships in the following decade, we should see an even distribution of clusters' "surviving" members across the clusters of the subsequent decade. For example, one of every seven members (14%) of Cluster 20A would proceed to each of the seven clusters in the thirties, one of every six members (17%) of Cluster 30A would be found in each cluster in the forties, and so on. For overall patterns to be established, the proportion of members of one cluster proceeding to another must be greater than expected by chance.[22]

(FIGURE 5 ABOUT HERE)

Indeed, some strong pathways emerge from these analysis (see Figure 5). In this representation, dark arrows indicate that at least twenty percent of one cluster's "surviving" members proceed to the same cluster in the subsequent decade (arrow style indicates the proportion). Grey arrows indicate the proportion of members of each cluster who do not proceed to the subsequent decade. Strong links between clusters are especially evident between the thirties and forties decades. Each thirties cluster sends at least forty percent of its members to a single cluster in the forties, and five of the seven send more than seventy percent of their members to one or two destination clusters. Ties between the twenties and thirties clusters are not as strong; still, four of the five twenties clusters send more than forty percent of their members to one or two clusters. Similarly, four of the six forties clusters send more than forty percent of their members to one or two clusters in the fifties. The overall pattern suggests both variability in the connections between life patterns in the twenties and life patterns in the thirties, and continuity over midlife, in which patterns established in the thirties tend to determine roles (and overlaps) over the following decade as well. This is consistent with an understanding of the life course in which timing, ordering, and even existence of family and career formation transitions vary in early adulthood, but the decisions made in this period – such as marriages, children, and possibly long-term employment – continue into later decades of life.

Specific pathways through clusters are anticipated by the similarities between clusters noted earlier. The pathway 20B–30F–40B–50D corresponds to the four female-dominated clusters in which part-time work, early childbearing, and low education and earnings are prevalent. The first three clusters in the similarly strong pathway 20D–30G–40D–50E disproportionately represent the older cohorts, and are characterized by low education, early family formation, and large families. Members of Cluster 20C are high achievers with later family formation, including many individuals with no children or only one child. The theme of no or few children continues through the pathway of 20C–30C–40F–50C, although income and education attainment cease to be clear differentiating factors. Although many members of these clusters have some children, it is notable that only two of the 151 individuals who have no children deviate from this pattern. Clearly, number of children is a key factor in classifying distances between life course graphs. Diagrams of life history charts for examples of these three pathways, drawn from the CCS, are show in Figures 6a, 6b, and 6c.

(FIGURES 6A, 6B, AND 6C ABOUT HERE)

Other paths are not as strongly patterned, but are still noteworthy. The seven predominantly male clusters partition into two rather fluid pathways. Both begin with Cluster 20A. Afterwards, portions of the pathway 20A–30A–40A–50E tends to be taken by men with higher education, very high earnings, and rather larger families, while portions of the pathway 20A–30E–40C–50A are taken by men with moderate education and earnings and families with two children. However, there is considerable bleed into and out of these pathways, particularly in the first, higher earnings path. This is largely due to cohort attrition, as more than half the individuals in Cluster 40A do not proceed to their fifties. On the other hand, the few men in this sample who are old enough to follow the last three legs of this pathway did not start in the same place: of the eight, none began with Cluster 20A (five began with Cluster 20E instead). A similar cohort relationship is found for the other pathway, in which only one of the twenty-three men followed the pathway 30E–40C–50A began with Cluster 20A (the majority began with Cluster 20C instead). Most of the movement in and out of these pathways involves clusters that are not male-dominated. Where there is cross-over between these two pathways, it is decidedly unidirectional. At both the transitions from twenties to thirties and thirties to forties, over twenty percent of men on the higher earnings pathway "defect" to the moderate earnings pathway, but only six and five percent, respectively, switch paths in the other direction. Unfortunately, space considerations constrain a more thorough examination of these findings. Diagrams of life history charts for examples of the pathway 20A–30A–40A–50B and 20C–30E–40C–50A are shown in Figures 6d and 6e.

(FIGURES 6D AND 6E ABOUT HERE)

While some of the pathways discussed here are relatively strong in terms of paired decade clusters, it is important to note that few individuals complete a "perfect" prototypical pathway through all four decades. To a large extent, their ability to do so is hampered by age: less than one-third of the sample (31%) have a cluster membership for the fifties decade. Only twenty women complete the partial prototypical female pathway 20B–30F–40B, and only four of these women are also found in 50D, along with one man who also completes the full pathway. However, those twenty-one people in 40B represent 60% of the survivors from 20B-30F, and the five people in 50D represent 38% of the people surviving after that point. This suggests that a larger sample of older individuals might substantiate the strength of these pathways. At the same time, preliminary exploration of the frequencies of particular pathways indicates substantial variation even among those with valid data for at least three of the four decade clusters. Unfortunately, a full examination of these findings and their implications is beyond the scope of this paper.

PROSPECTS

As we have shown, the interval graph approach to the representation and analysis of life history data is uniquely suited to addressing a variety of questions regarding the life course. In conceptualizing the life course as a pattern of intersecting roles, we draw on key themes in life course research, bringing together the context of concurrent roles with the accumulated context of experiences over time (e.g., Elder 1995; Moen, Dempster-McClain, and Williams 1992; Moen and Wethington 1999; Waldron, Weiss, and Hughes 1998). Analyzing overlaps between spells in addition to current status and number of past spells can add substantially to the understanding of different life course experiences. Such a focus does not preclude the inclusion of other indicators of life course patterns: timing measures (e.g., age at birth of first child) and sequence measures (e.g., time before or after first marriage that the first full-time job was started) can be used to supplement comparisons between patterns. Furthermore, the approach is highly adaptable to specific research foci: a variety of overlap measures can be used, for any set of roles, over recent or long-term time frames. Thus, it encourages asking questions that previously would have been difficult or impossible to answer.

Descriptive characteristics of life course graphs can be used to address a wide range of life course issues. For instance, to what extent do overlaps between family roles and professional versus nonprofessional jobs in the twenties and thirties help explain later occupational outcomes? Does the proportion of time in which child spells overlap with parents' full-time jobs or part-time jobs have implications for child or parent well-being? Which patterns of overlaps between elder care spells, child care spells, and work spells are related to lower stress levels for those giving and receiving the care? Do people with long histories of densely connected roles display different strategies for negotiating their multiple roles than people who have only recently taken on new roles? Is the reaction to a loss of a role affected by the extent to which that role has tied together a person's life epoch up to that point, or by whether it ends that epoch? Does the separation of lives into disconnected epochs have particular significance for people who are struggling to leave behind abusive relationships, deviance careers, or substance addiction? Framing questions in terms of overlap patterns is clearly consistent with life course researchers' focus on context and multiple roles over time.

The interval graph approach also introduces new ways of comparing life course patterns *in toto*, by calculating interstructural (i.e., interpersonal) distances between life course graphs based either on number of overlaps or lengths of overlaps. In this paper, we have illustrated the use of those distances to identify clusters of patterns within each decade of the life course. This demonstration hints at the potential utility of this method. One possible application would begin with determining interstructural distances in large, representative samples. Multivariate analyses such as MDS and canonical correlation can test the extent to which distance scores from alternate model specifications map onto theoretically important dimensions of difference. Cluster analyses would be a natural next step. If our assumption of distinct types of life trajectories holds true, and our distance specifications accurately model those assumptions, key clusters should repeatedly emerge that represent those common pathways. Once established, descriptions of these key clusters could be used to classify individuals' life course patterns in smaller data sets which include life history data but for which distance scores and cluster analyses would not be feasible.

Other applications might focus on predicting later life course patterns using early indicators. This is an extension of identifying pathways through decades by cluster, in which we ask the extent to which individuals' later overlap patterns can be predicted using overlap patterns in early adulthood. Alternately, distance scores can be used to assess within- and between-group effects on differences in life course patterns. For example, are the life histories of female professionals more similar to those of female non-professionals or to those of their male colleagues?

In addition, the interval graph approach is well-suited to the examination of linked lives. A simple example of such an analysis involves comparing spouses on graph-level indices such as the number of overlaps between jobs and child co-residence spells. At the interstructural level, dyadic distances indicate how different spouses' life patterns have been. Distances between spouses' overlap patterns may reveal much about how certain types of similarities and differences contribute to later outcomes such as marital support or conflict or disparities in wives' and husbands' occupational attainment. Differences may be indicative of couple's strategies: spouses who report nontraditional attitudes should exhibit more similar life course patterns than those with traditional attitudes. Given the flexibility of the interval graph approach, a researcher could include spells from both spouses in one graph that represents the "linked life course" of their relationship. This would produce within-spouse and between-spouse overlaps: not only between his part-time job and his educational spell, but also between his educational spell and her full-time job.

In summary, we argue that the interval graph approach to studying life course patterns has great potential for addressing a wide range of life course questions. The relational nature of the associated techniques, originally developed for interpersonal networks, lends itself well to the study of the life course as an intrapersonal network of intersecting roles across time. As such, we believe it to be a useful addition to the current repertoire of methods.

REFERENCES

Abbott, Andrew and Alexandra Hrycak. 1990. "Measuring Resemblance in Sequence Data: An Optimal Matching Analysis of Musicians' Careers." *American Journal of Sociology* 96:144-85.

Abbott, Andrew and Angela Tsay. 2000. "Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect." *Sociological Methods and Research* 29:3-33.

Anderson, Brigham S., Carter T. Butts, and Kathleen M. Carley. 1999. "The Interaction of Size and Density with Graph-Level Indices." *Social Networks* 21:239-267.

Banks, David L. and Kathleen M. Carley. 1994. "Metric Inference for Social Networks." *Journal of Classification* 11:121-149.

Blair-Loy, Mary. 1999. "Career Patterns of Executive Women in Finance: An Optimal Matching Analysis." *The American Journal of Sociology* 104:1346-1397.

Butts, Carter T. 1998. "Cluster Analysis of Unlabeled Structures." ICES Research Report, #88-04-98, Carnegie Mellon University, Pittsburgh, PA.

Butts, Carter T. and Kathleen M. Carley. 1998. "Canonical Labeling to Facilitate Graph Comparison." ICES Research Report, #88-06-98, Carnegie Mellon University, Pittsburgh, PA.

Butts, Carter T. and Kathleen M. Carley. 2001. "Multivariate Methods for Interstructural Analysis." CASOS Working Paper, Carnegie Mellon University, Pittsburgh, PA.

Butts, Carter T. and Joy E. Pixley. 2001. "An Interval Graph Approach to the Representation of Life History Data." CASOS Working Paper, Carnegie Mellon University, Pittsburgh, PA.

Drobnic, Sonja, Hans-Peter Blossfeld, and Gotz Rohwer. 1999. "Dynamics of Women's Employment Patterns over the Family Life Course: A Comparison of the United States and Germany." *Journal of Marriage and the Family* 61:133-146.

Elder, Glen H. 1995. "The Life Course Paradigm: Social Change and Individual Development." Pp. 101-139 in *Examining Lives in Context: Perspectives on the Ecology of Human Development*, edited by P. Moen, G. H. Elder, Jr., and K. Lüscher. Washington DC: American Psychological Association.

Elder, Glen H., Jr., Michael J. Shanahan, and Elizabeth Colerick Clipp. 1994. "When War Comes to Men's Lives: Life-Course Patterns in Family, Work, and Health." *Psychology and Aging* 9:5-16.

George, Linda K. 1993. "Sociological Perspectives on Life Transitions." *Annual Review of Sociology* 19:353-373.

Giele, Janet Z. 1998. "Innovation in the Typical Life Course." Pp. 231-263 in *Methods of Life Course Research: Qualitative and Quantitative Approaches*, edited by J. Z. Giele and G. H. Elder, Jr. Thousand Oaks, CA: Sage.

Giele, Janet Z. and Glen H. Elder, Jr. 1998. "Life Course Research: Development of a Field." Pp. 5-27 in *Methods of Life Course Research: Qualitative and Quantitative Approaches*, edited by J. Z. Giele and G. H. Elder, Jr. Thousand Oaks, CA: Sage.

Hamming, R. W. 1950. "Error Detecting and Error Correcting Codes." *Bell System Technical Journal* 29:147-160.

Han, Shin-Kap and Phyllis Moen. 1999. "Work and Family Over Time: A Life Course Approach." *Annals of the American Academy of Political and Social Science* 562:98-110.

Hotchkiss, Julie L. and Robert E. Moore. 1999. "On the Evidence of a Working Spouse Penalty in the Managerial Labor Market." *Industrial and Labor Relations Review* 52:410-423.

Levine, Joel H. 2000. "But What Have You Done for Us Lately?: Commentary on Abbott and Tsay." *Sociological Methods and Research* 29:34-40.

Long, J. Scott, Paul D. Allison, and Robert McGinnis. 1993. "Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity." *American Sociological Review* 58:703-722.

Moen, Phyllis. 1997. "Women's Roles and Resilience: Trajectories of Advantage or Turning Points?" Pp. 133-56 in *Stress and Adversity Over the Life Course: Trajectories and Turning Points*, edited by I. H. Gotlib and B. Wheaton. New York: Cambridge University Press.

Moen, Phyllis, Donna Dempster-McClain, and Robin M. Williams, Jr. 1992. "Successful Aging: A Life-Course Perspective on Women's Multiple Roles and Health." *American Journal of Sociology* 97:1612-1638.

Moen, Phyllis, Glen Elder, Jr., and Kurt Luscher. 1995. *Examining Lives in Context: Perspectives on the Ecology of Human Development*. Washington, D.C.: American Psychological Association.

Moen, Phyllis and Elaine Wethington. 1999. "Midlife Development in a Life Course Context." in *Life in the Middle: Psychological and Social Development in Middle Age*, edited by S. L. Willis and J. D. Reid. San Diego: Academic Press.

Shanahan, Michael J. 2000. "Pathways to Adulthood in Changing Societies: Variability and Mechanisms in Life Course Perspective." *Annual Review of Sociology* 26:667-692.

Singer, Burton, Carol D. Ryff, Deborah Carr, and William J. Magee. 1998. "Linking Life Histories and Mental Health: A Person-Centered Strategy." *Sociological Methodology* 28:1-51.

Spain, Daphne and Suzanne M. Bianchi. 1996. *Balancing Act: Motherhood, Marriage, and Employment among American Women*. New York: Russell Sage Foundation.

Waldron, Ingrid, Christopher Weiss, and Mary Elizabeth Hughes. 1998. "Interacting Effects of Multiple Roles on Women's Health." *Journal of Health and Social Behavior* 39:216-236.

Wasserman, Stanley. 1980. "Analyzing Social Networks as Stochastic Processes." *Journal of the American Statistical Association* 75:280-294.

Wasserman, Stanley and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. New York, NY: Cambridge University Press.

Wasserman, Stanley and S. O. Weaver. 1985. "Statistical Analysis of Binary Relational Data: Parameter Estimation." *Journal of Mathematical Psychology* 29:406-427.

Wu, Lawrence L. 2000. "Some Comments on "Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect"." *Sociological Methods and Research* 29:41-64.

Wu, Lawrence L. and Brian C. Martinson. 1993. "Family Structure and the Risk of a Premarital Birth." *American Sociological Review* 58:210-232.

ENDNOTES

[1] For instance, it is expected that stochastic network models (such as the p^* family of loglinear models (Wasserman 1980; Wasserman and Weaver 1985)) will prove applicable to predicting life course patterns (e.g., modeling probabilities of inter-domain overlaps in terms of demographic characteristics and prior experiences).

[2] Earlier in the same discussion, Abbot explained: "A number of authors have created complex 'events' for their sequences by cross-classifying a number of simple events. This is the only strategy available for dealing with multiple, parallel tracks of sequence information in the OM [optimal matching] framework; it must be reduced somehow to the unilinear structure expected by the OM algorithms. … These various combinatoric codings of events create serious complexities in the setting of replacement costs..." (Abbott and Tsay 2000: 9-10).

[3] It is also possible to construct spell-level indices, such as the centrality of particular marital spell vis-à-vis all other spells in the graph (see Butts and Pixley 2001). Such indices are better suited to in-depth case studies of individual lives than to broader comparisons between life course patterns.

[4] It is important, however, to take into account the distributional properties of these indices when deploying them; see Anderson et al. (1999) for a discussion of these issues.

[5] The definition of structural distances for partially labeled structures is closely related to the optimal matching methods of sequence analysis; the former can be regarded as resulting from the latter as applied to the graph's accessible permutation group (Butts and Carley, 2001).

[6] More precisely, the betweenness centrality score is moderated when multiple geodesics exist between pairs of vertices, such that a vertex on a redundant path obtains a betweenness score proportionally lower depending upon how many other paths of the same (shortest) length exist.

[7] For example, while v_c is linked to both v_d and v_f, there is a direct (and thus shorter) path between v_d and v_f.

[8] A star formation, where one actor is tied to all other actors, who are tied only to the first actor, has the highest score (1.0) for all three types of centralization.

[9] When a graph is disconnected (when there exist multiple components), indices based on the distance between points (e.g., mean geodesic distance, diameter, and closeness) are not well-defined. A few approaches exist, none of which are ideal. The choice made here is to assume, when two vertices are not connected by edges, that the distance between them is the maximum possible, or $n-1$. A healthy skepticism about the utility of these indices should thus be employed when used with disconnected graphs.

[10] While the intervals listed here are all temporal, intervals may conceivably be spatial as well. For example, an interval graph could be used to describe the extent to which companies' sales regions overlap.

[11] While continuity often carries positive valence, scholars interested in "career" pathways of substance abuse, mental illness, or criminality should note that epochs represent continuity in relationships with people – partners, children, and coworkers – who may be contributing to the behavior or condition under study.

[12] In classical network analysis, block modeling groups vertices (e.g., people) based on a criterion such as structural equivalence, allowing examination of ties within blocks and between different blocks. Here, spells within the same life domain (e.g., full-time jobs, or children) are considered to be equivalent.

[13] Managing vertex exchangeability and addition of spells so that both graphs have the same number of each type of spell can be accomplished in a number of ways. While the statistical properties of the primary options have been determined, additional work is being done to match methods to theoretical bases for such decisions.

[14] For example, metrics based on sums of squares prioritize minimizing larger distances over minimizing the equivalent sum of smaller distances. By contrast, absolute difference gives equal priority to all distances and is more robust to outliers.

[15] So, while Hamming distance would count the number of overlaps that did not match, absolute distance would produce the same score by summing the differences of unmatched overlaps of "length" one and "length" zero.

[16] It is also possible to create indices that reflect tie values. For instance, a generalized measure of degree centrality can be employed that incorporates not only the number of overlaps with other spells, but also the amount of the overlap. See Wasserman and Faust (1994) for some examples of structural indices on valued graphs.

[17] These Census tracts were selected based on their representation in an earlier study, also conducted by the Cornell Employment and Family Careers Institute, such that the more representative households in the CCS could be compared to the select sample of highly-educated, middle-class dual-earner couples used in the first survey. Urban, suburban, and semi-rural areas are included.

[18] As pointed out earlier, the geodesic (or shortest path between two vertices) is not well-defined for graphs with multiple components, in which some vertex pairs are not connected. For such graphs, the approach used here is to define the geodesic between two unconnected vertices to be one hop greater than the longest possible path: that is, one encompassing all vertices in the graph. Hence, for disconnected graphs, the diameter (as the longest geodesic) is defined by the number of vertices.

[19] The relationship between age and income in this sample is neither linear nor quadratic; age categories provided the best fit to the data.

[20] Graphs (or sub-graphs) with fewer than two spells are by definition degenerate, and are omitted from these calculations. Because of this, we are missing a cluster membership (and in one case, two) for twenty-three people. In the twenties, thirties, and forties, this group consists primarily of people who had one full-time job, with no marital, child, or educational spells during that decade. In the fifties cluster, nine women and two men are married but report no other spells.

[21] Cluster analyses presented here were performed using Ward's method. This algorithm was judged to produce clusters most consistent with a visual inspection of MDS (multidimensional scaling) plots. Similar results are obtained using other clustering algorithms.

[22] An alternate approach might focus on the size of the destination clusters, arguing that we would expect larger proportions of source clusters to proceed to larger clusters. For instance, if fifty percent of individuals in their forties were found in a certain cluster, it should not be surprising if around half of each cluster in the thirties led to this cluster. This approach begs the question: the reason the hypothetical forties cluster is so large is *because* such a large proportion of many smaller clusters all led to it, and certainly such a result would indicate a clear pathway effect. In any case, membership in clusters in the thirties, forties, or fifties are relatively evenly distributed. The largest cluster in the thirties has 22% of the age group (versus an estimated 14% each), the largest cluster in the forties has 21% of the age group (versus an estimated 16% each), and the largest cluster in the fifties has 24% of the age group (versus an estimated 20% each). Readers who prefer the alternate argument should thus ignore any pathways marked as being in the 20%-29% range.

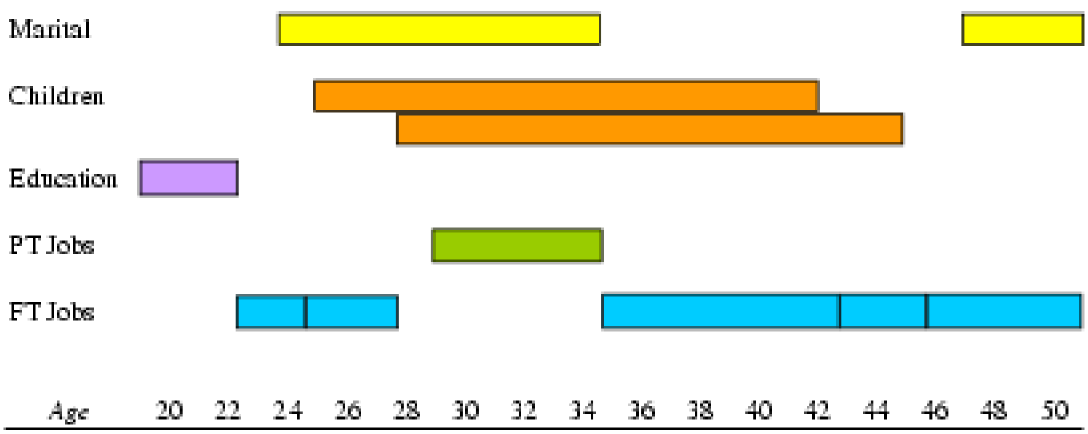

Figure 1. Example of a Life Course Chart

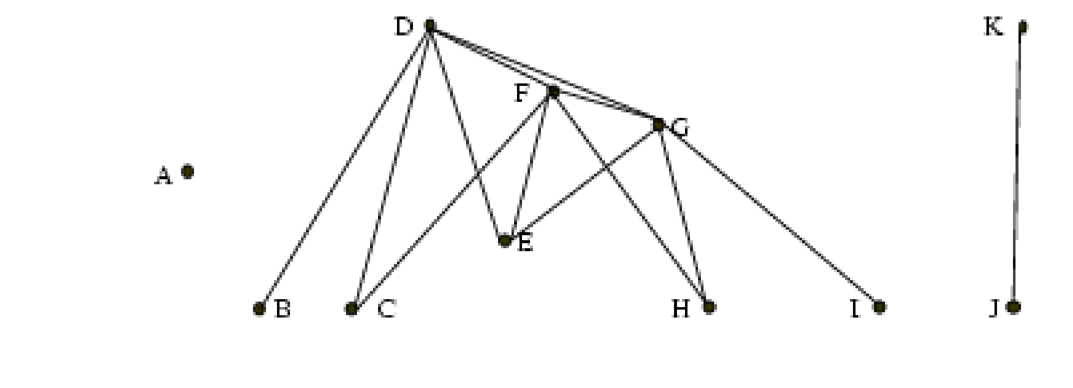

Figure 2. Example of a Simple Graph

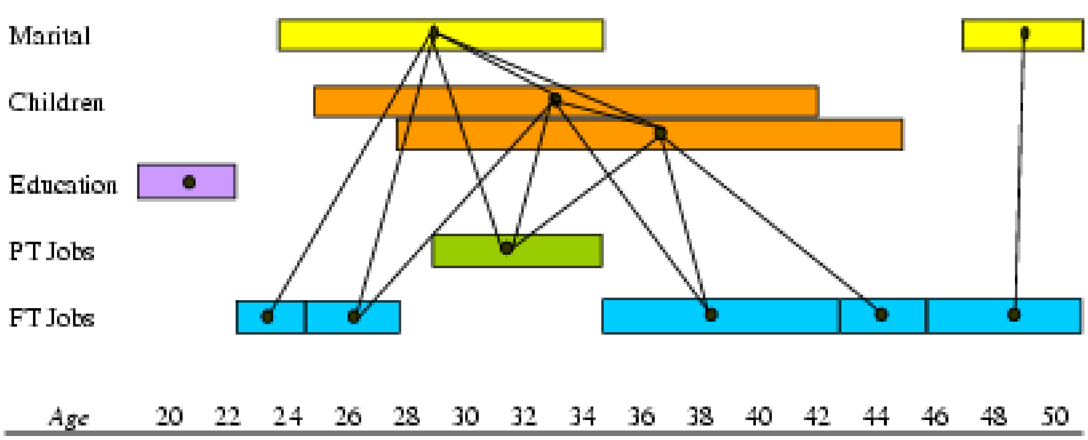

Figure 3. Representation of a Simple Life History Interval Graph

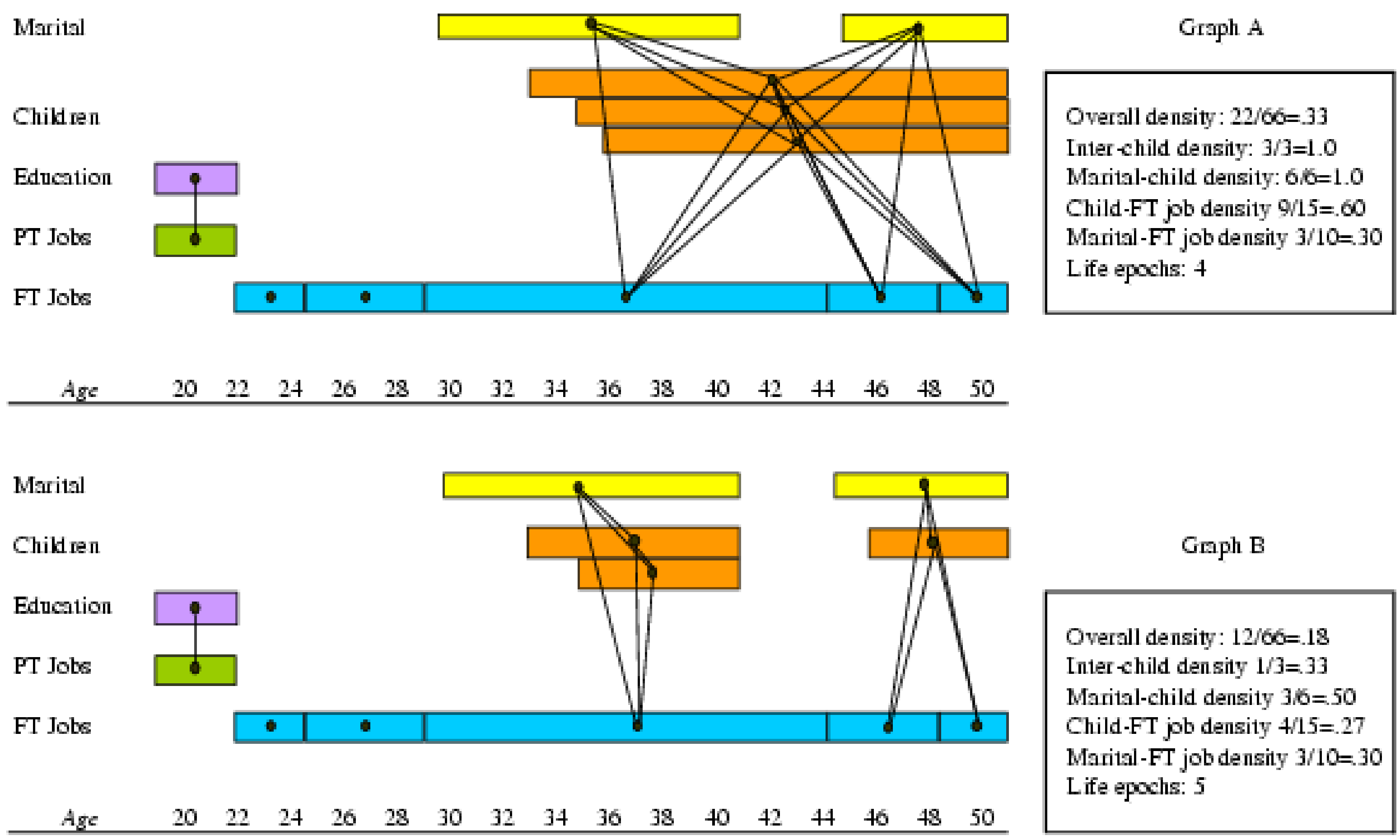

Figure 4. Comparing Densities Between Two Life Graphs with the Same Number of Spells.

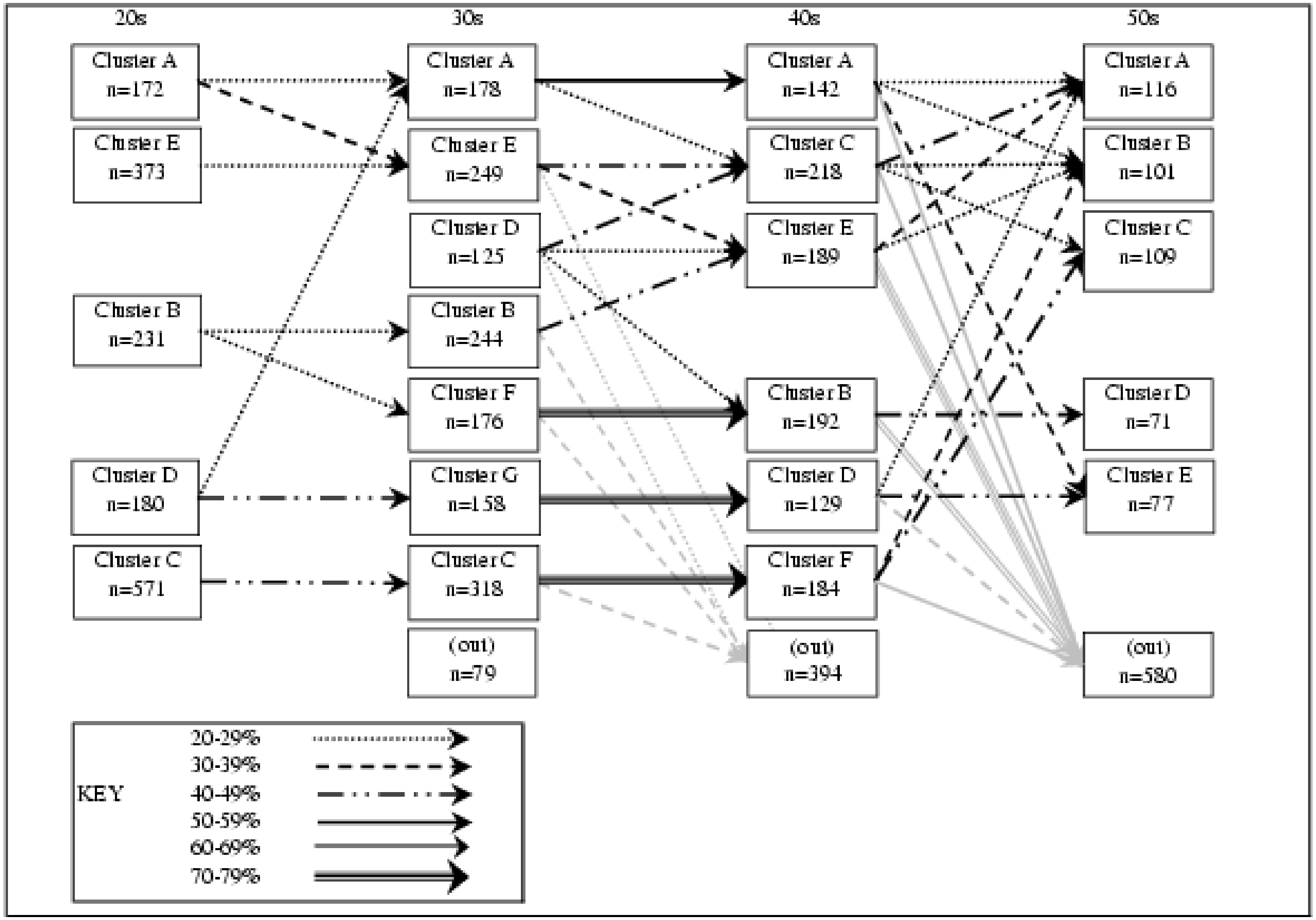

Figure 5. Diagram of the Strongest Pathways between Decade Clusters

Source: Cornell Community Study (1999-2000); 1,515 married and 17 single men and women in upstate New York.
Note: Grey lines represent the proportion of each cluster that has no data for the next decade of life. Dark lines represent the proportion of those remaining who proceed to particular clusters in the next decade.

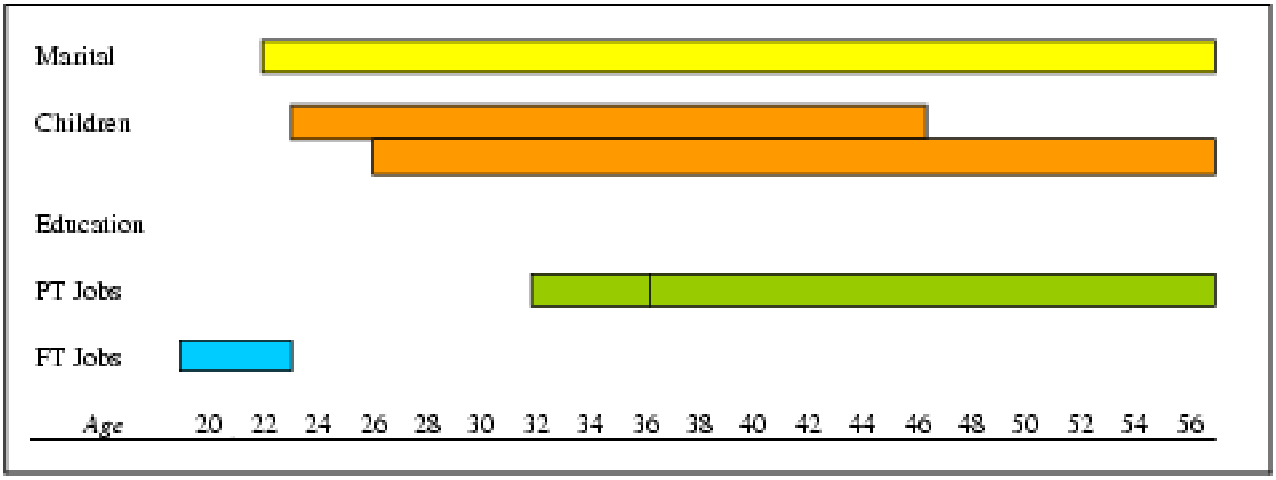

Figure 6a. Example of Pathway 20B - 30F - 40B - 50D.

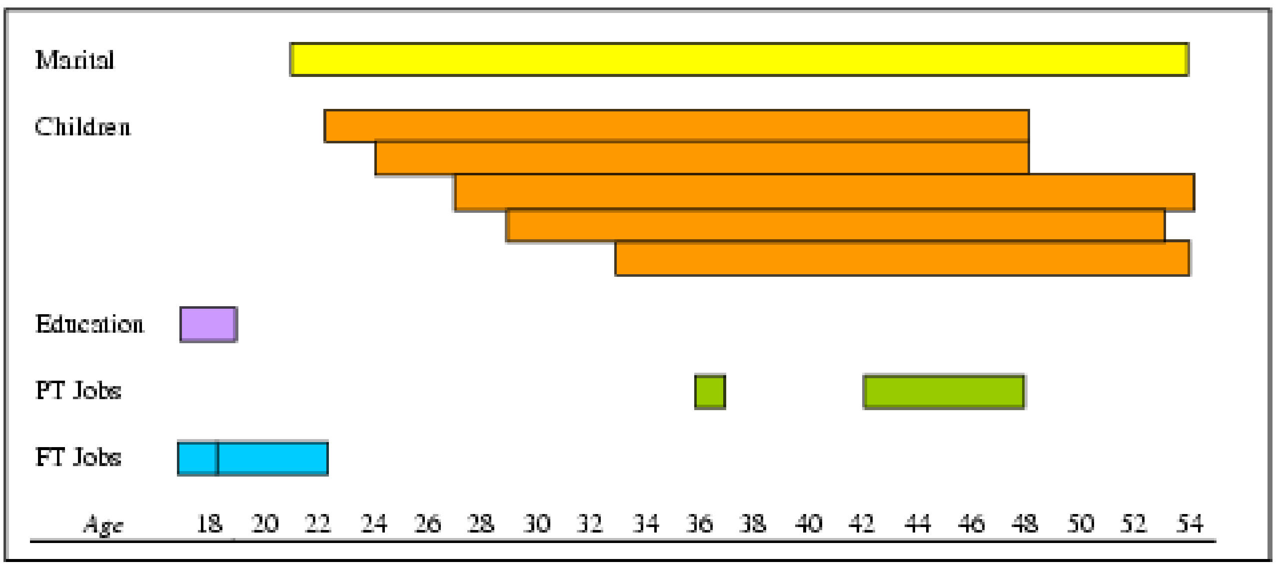

Figure 6b. Example of Pathway 20D - 30G - 40D - 50E.

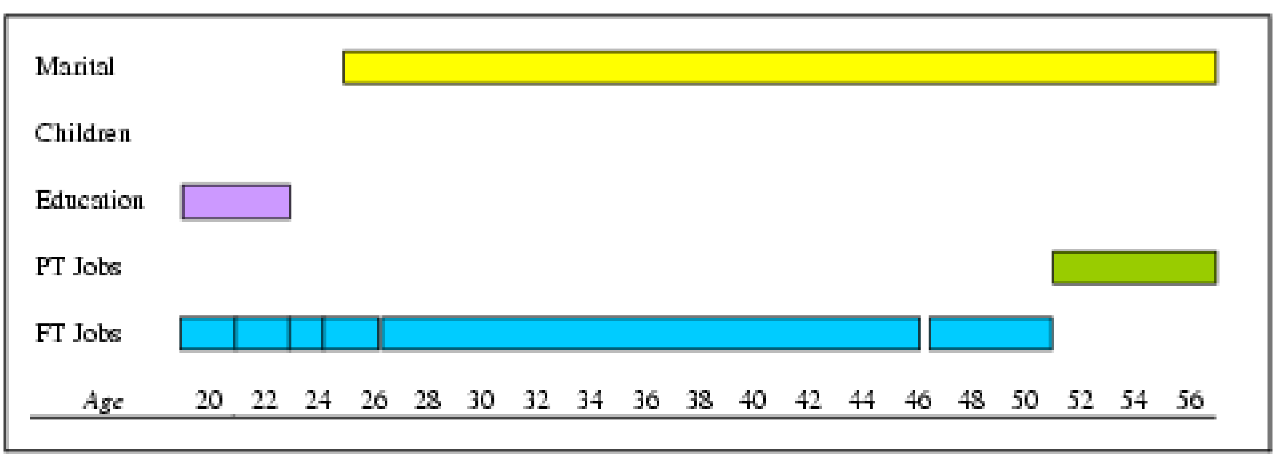

Figure 6c. Example of Pathway 20C - 30C - 40F - 50C.

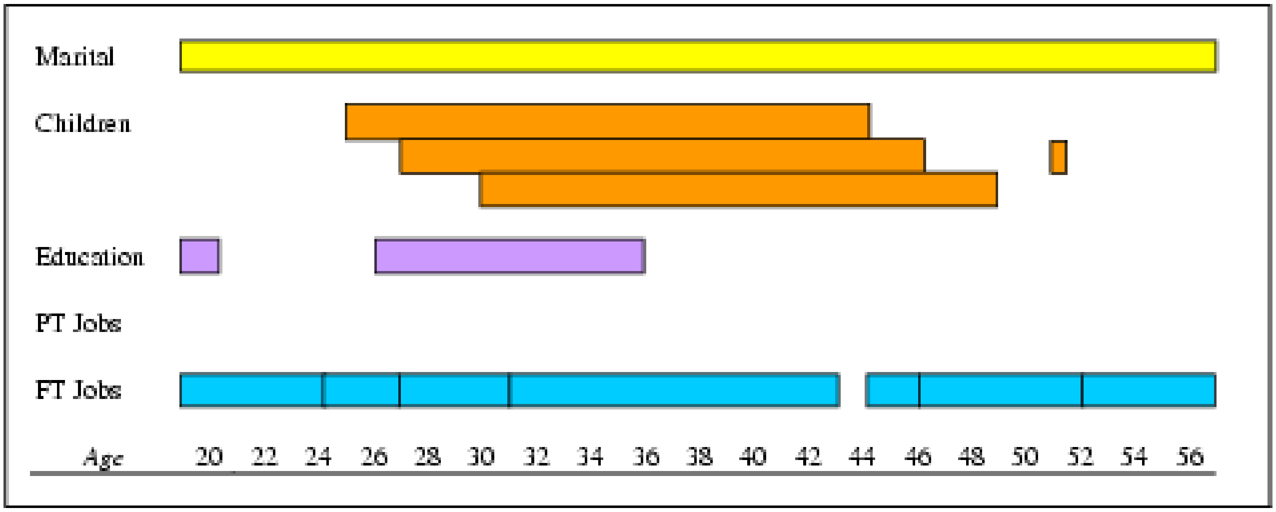

Figure 6d. Example of Pathway 20A - 30A - 40A - 50B.

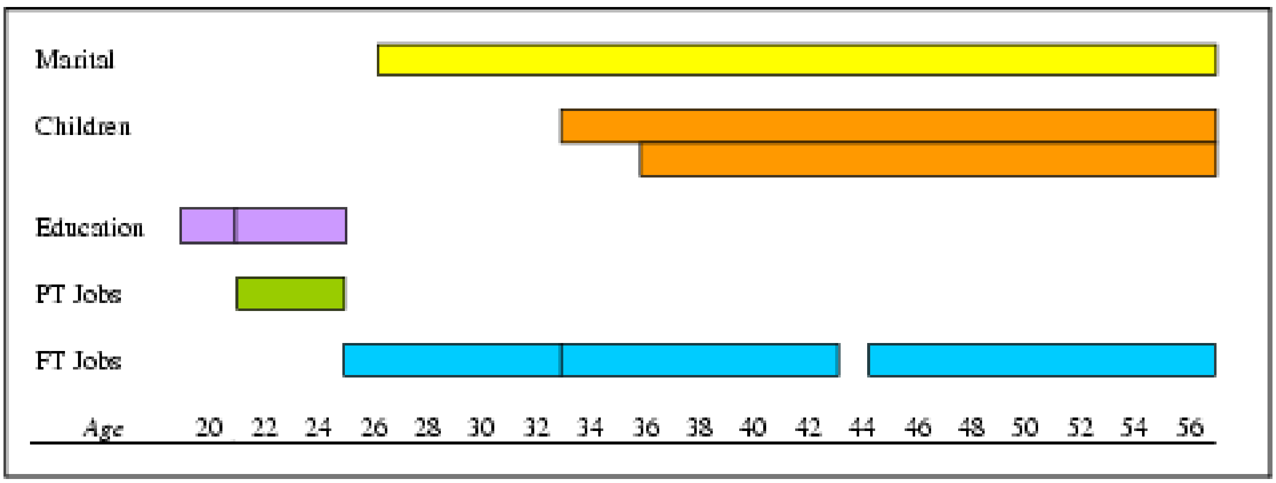

Figure 6e. Example of Pathway 20C - 30E - 40C - 50A.

Table 1. Descriptions of Hypothetical Life History Examples

	Marr.	Div.	Marital Timing[1]	Children	PT jobs	FT jobs	Worked with children	Educ. spells
Cluster 1								
Intermittent working parent [2]	1	0	early	2	3	2	PT	0
Intermittent working delayer parent [2]	1	0	late	2	4	2	PT	2
Intermittent working remarried parent [2]	2	1	early	2	2	3	PT + FT	0
Cluster 2								
Multiple relationship delayer nonparent	3	3	late	0	0	6	---	1
Multiple relationship delayer nonparent w/PT work	3	3	late	0	2	5	---	1
Steady working delayer nonparent	1	0	late	0	0	3	---	1
Cluster 3								
Working remarried delayer parent	2	1	late	2	1	5	FT	2
Working delayer with second family	2	1	late	3	1	5	FT	2
Cluster 4								
Married working parent	1	0	early	2	0	6	FT	0
Married working delayer parent	1	0	late	2	1	5	FT	1
Single working parent	1	1	early	2	1	5	FT	0
Single working parent with later education	1	1	early	2	0	4	FT	2

Note: Distance between members within clusters indicates groups that splintered in the next level of clustering.

[1] For parents, the first child followed within two years of first marriage.

[2] Intermittent working parents have periods of not working during their child spells.

Table 2. Correlation between Age and Life History Graph Indices

	r	*p*	N		*r*	*p*	N
Standard Graph Indices				*Intra-domain overlap density* [2]			
Number of spells	.39	<.0001	1532	FT employment	-.02	.4154	1371
Number of overlaps	.39	<.0001	1532	PT employment	-.09	.1184	296
Number of epochs	-.04	.1320	1532	Marital	.09	.1089	323
Median spells per epoch	.20	<.0001	1532	Education	.03	.4982	703
Density	.01	.6073	1532	Children	-.32	<.0001	1202
Median degree	.26	<.0001	1532				
Diameter	.20	<.0001	1531	*Inter-domain overlap density* [3]			
Mean geodesic distance	.07	.0068	1531	FT and PT employment	-.08	.0407	644
Degree Centralization	.09	.0003	1532	FT and Marital	.04	.1286	1503
Betweenness Centralization	.08	.0010	1532	FT and Education	-.02	.4530	1284
Closeness Centralization	.01	.6497	1532	FT and Children	.01	.7020	1364
				PT and Marital	.11	.0054	653
Number of spells of each type				PT and Education	-.20	<.0001	556
FT employment	.24	<.0001	1532	PT and Children	-.09	.0290	602
PT employment	.04	.1077	1532	Marital and Education	.11	<.0001	1288
Marital	.13	<.0001	1532	Marital and Children	-.11	<.0001	1380
Education	.08	.0011	1532	Education and Children	.02	.4130	1166
Children [1]	.39	<.0001	1532				

Source: Cornell Community Study (1999-2000); 1,515 married and 17 single men and women in upsta

** $p < .01$; * $p < .05$; + $p < .10$

[1] Children spells refer to periods of co-residence, and include biological and non-biological children.

[2] Overlap density is only meaningful for cases with two or more spells in the same category; other cases

[3] Overlap density is only meaningful for cases with at least one spell of each type; other cases are missi

Table 3. Graph-Level Indices by Gender

	Men			Women			
	N	Mean	SD	N	Mean	SD	p[a]
Standard Graph Indices							
Number of spells	760	10.96	(4.00)	772	10.12	(3.55)	<.0001
Number of overlaps	760	51.92	#####	772	44.35	(26.58)	<.0001
Number of epochs	760	2.56	(1.50)	772	2.43	(1.42)	.0731
Median spells per epoch	760	4.97	(4.30)	772	5.03	(4.18)	.4605
Density	760	.37	(.16)	772	.37	(.16)	.9597
Median degree	760	3.30	(1.59)	772	3.07	(1.52)	.0025
Mean geodesic distance	760	3.96	(2.59)	772	3.63	(2.29)	.0195
Diameter	760	8.96	(5.11)	772	7.96	(4.55)	.0002
Degree Centralization	759	.47	(.17)	772	.49	(.17)	.0594
Betweenness Centralization	759	.26	(.20)	772	.29	(.20)	.0007
Closeness Centralization	760	.15	(.13)	772	.16	(.13)	.0160
Number of spells in each domain							
FT employment	760	5.3	(2.9)	772	3.9	(2.5)	<.0001
PT employment	760	.4	(.7)	772	1.2	(1.4)	<.0001
Marital	760	1.3	(.6)	772	1.2	(.5)	.8482
Education	760	1.6	(1.2)	772	1.5	(1.0)	.0772
Child	760	2.5	(1.6)	772	2.4	(1.5)	.3816
Number of intra-domain overlaps							
FT employment	760	.249	(.917)	772	.150	(.535)	.0102
PT employment	760	.030	(.225)	772	.101	(.444)	<.0001
Education	760	.068	(.282)	772	.053	(.236)	.3260
Child	760	2.393	#####	772	2.260	(3.274)	.9346
Intra-domain overlap density[b]							
FT employment	712	.02	(.08)	659	.04	(.16)	.1334
PT employment	57	.21	(.38)	239	.11	(.27)	.1408
Education	358	.07	(.23)	345	.07	(.22)	.4769
Child	595	.90	(.23)	607	.93	(.19)	.0239

	Men			Women			
	N	Mean	SD	N	Mean	SD	p[a]
Number of inter-domain overlaps							
FT and PT employmen	760	.22	(.73)	772	.18	(.58)	.9946
FT and Marital	760	4.36	(2.75)	772	3.01	(2.32)	<.0001
FT and Education	760	1.23	(1.73)	772	.84	(1.28)	<.0001
FT and Child	760	7.01	(6.35)	772	3.44	(4.42)	<.0001
PT and Marital	760	.27	(.67)	772	1.06	(1.38)	<.0001
PT and Education	760	.12	(.46)	772	.24	(.66)	<.0001
PT and Child	760	.39	(1.27)	772	1.92	(2.97)	<.0001
Marital and Education	760	.73	(.92)	772	.63	(.92)	.0141
Marital and Child	760	2.59	(1.90)	772	2.56	(1.71)	.6936
Child and Education	760	.83	(1.60)	772	.68	(1.47)	.0231
Inter-domain overlap density[c]							
FT and PT employmen	192	.18	(.28)	452	.06	(.16)	<.0001
FT and Marital	751	.73	(.27)	752	.69	(.30)	.0102
FT and Education	651	.17	(.21)	633	.16	(.22)	.0625
FT and Child	684	.56	(.23)	680	.39	(.30)	<.0001
PT and Marital	192	.63	(.42)	461	.79	(.33)	<.0001
PT and Education	162	.23	(.33)	394	.14	(.25)	.0080
PT and Child	173	.43	(.42)	429	.65	(.37)	<.0001
Marital and Education	647	.36	(.38)	641	.32	(.37)	.0168
Marital and Child	686	.89	(.21)	694	.92	(.18)	.0799
Child and Education	591	.19	(.28)	575	.17	(.28)	.1074

Source: Cornell Community Study (1999-2000); 1,515 married and 17 single men and women in upstate New York.

[a] Indicates p-level for gender difference (Kruskal-Wallis test).

[b] Overlap density is only meaningful for cases with two or more spells in the same category; other cases are missing.

[d] Overlap density is only meaningful for cases with at least one spell of each type; other cases are missing.

Table 4. Regression of Income (Logged) on Standard Predictors and Graph-Level Indices.

	Women			Men		
	b	se	p	b	se	p
Intercept	8.6538	.1419	<.0001	9.8910	.1312	<.0001
Bachelor's degree	.1991	.0641	.0020	.2783	.0528	<.0001
Advanced degree	.6508	.1595	<.0001	.6861	.1014	<.0001
Work hours	.0283	.0022	<.0001	.0115	.0017	<.0001
Age 30-39	.2398	.1218	.0494	.2610	.1142	.0226
Age 40-49	.3547	.1206	.0034	.3981	.1127	.0004
Age 50-59	.3387	.1298	.0093	.3047	.1179	.0100
Age 60 and older	.4505	.1722	.0091	.1224	.1357	.3674
Number of part-time job spells	-.0332	.0212	.1172	-.1456	.0320	<.0001
Number of child spells	.1059	.0302	.0005	-.0003	.0236	.9899
Number of educational spells	-.0968	.0250	.0001	-.0144	.0211	.4945
Full-time job--child overlaps	.0282	.0075	.0002	.0065	.0052	.2152
N	588			705		
R^2	.44			.25		
Adjusted R^2	.43			.24		
F value	40.54	<.0001		20.93	<.0001	
Standard error	.65			.60		

Source: Cornell Community Study (1999-2000).

Table 5. Current Characteristics of Members of Clusters in Each Previous Decade of Life

	Total	CLUS A	CLUS B	CLUS C	CLUS D	CLUS E	CLUS F	CLUS G
Twenties								
N	1527	172	231	571	180	373		
age 20-29	5%	2%	3%	9%	2%	5%		
age 30-39	26%	23%	29%	26%	15%	31%		
age 40-49	37%	38%	36%	38%	34%	37%		
age 50-59	23%	26%	25%	20%	32%	22%		
60 or older	8%	11%	7%	6%	17%	6%		
Female	50%	26%	76%	46%	57%	50%		
Single	1%	1%	0%	2%	0%	1%		
Married	94%	96%	96%	91%	98%	94%		
Cohabiting	5%	3%	4%	6%	2%	5%		
Only one relationship	78%	73%	83%	77%	80%	79%		
Age at first cohab/marriage	24.4	22.2	22.3	27.4	21.3	23.9		
College degree	35%	33%	31%	37%	32%	38%		
Advanced degree	20%	15%	12%	29%	13%	17%		
Income	$49,353	$52,834	$33,541	$55,159	$46,123	$49,542		
Not working	16%	10%	20%	13%	21%	17%		
Part-time	14%	10%	20%	13%	13%	16%		
Full-time	70%	80%	60%	74%	67%	67%		
Age at first child	27.0	23.8	24.3	32.6	21.8	26.7		
No children	10%	0%	0%	26%	0%	0%		
One child	12%	0%	0%	19%	0%	22%		
Two children	40%	49%	55%	37%	1%	48%		
Three or more	38%	51%	45%	18%	99%	30%		
Thirties								
N	1448	178	244	318	125	249	176	158
age 30-39	27%	20%	32%	33%	22%	27%	30%	18%
age 40-49	39%	45%	37%	35%	39%	45%	42%	34%
age 50-59	25%	26%	23%	24%	30%	23%	24%	27%
60 or older	8%	9%	8%	7%	8%	6%	4%	22%
Female	50%	24%	59%	46%	54%	33%	93%	53%
Married (v. cohabiting)	95%	99%	98%	89%	97%	95%	98%	96%
Only one relationship	78%	85%	82%	71%	79%	75%	90%	70%
Age at first cohab/marriage	24.4	23.6	24.2	26.6	24.2	24.3	23.7	22.9
College degree	35%	40%	34%	31%	33%	37%	33%	35%
Advanced degree	20%	20%	12%	25%	43%	18%	16%	15%
Income	$50,416	$64,671	$51,006	$49,849	$52,032	$59,325	$25,514	$42,919
Not working	16%	6%	30%	15%	15%	6%	18%	25%
Part-time	14%	3%	9%	9%	14%	4%	56%	15%
Full-time	70%	91%	70%	76%	71%	90%	26%	59%
Age at first child	27.1	25.3	26.5	31.9	27.2	27.1	26.7	23.9
No children	7%	0%	1%	31%	0%	0%	0%	0%
One child	12%	0%	1%	52%	1%	1%	0%	0%
Two children	41%	0%	57%	12%	84%	86%	56%	0%
Three or more	40%	100%	41%	5%	15%	13%	44%	100%

Table 5, continued

	Total	CLUS A	CLUS B	CLUS C	CLUS D	CLUS E	CLUS F
Forties							
N	1054	142	192	218	129	189	184
age 40-49	54%	55%	57%	48%	37%	69%	54%
age 50-59	34%	32%	36%	39%	36%	23%	37%
60 or older	12%	11%	7%	14%	26%	7%	9%
Female	48%	27%	83%	33%	39%	55%	45%
Married (v. cohabiting)	96%	96%	96%	97%	98%	97%	89%
Only one relationship	76%	87%	82%	70%	74%	84%	63%
Age at first cohab/marriage	24.2	24.4	23.5	24.1	23.1	24.2	25.8
College degree	34%	43%	35%	35%	35%	32%	26%
Advanced degree	23%	24%	23%	27%	17%	18%	24%
Income	$52,652	$70,710	$31,810	$57,065	$50,853	$53,750	$54,306
Not working	16%	7%	13%	12%	21%	26%	15%
Part-time	14%	3%	53%	3%	9%	3%	9%
Full-time	71%	90%	34%	85%	71%	71%	77%
Age at first child	27.1	26.5	27.1	27.0	24.7	27.5	30.1
No children	5%	0%	0%	0%	0%	0%	28%
One child	10%	0%	1%	3%	0%	1%	55%
Two children	40%	3%	53%	65%	1%	80%	10%
Three or more	45%	97%	46%	32%	99%	20%	7%
Fifties							
N	474	116	101	109	71	77	
age 50-59	74%	72%	73%	77%	80%	69%	
60 or older	26%	28%	27%	22%	18%	31%	
Female	43%	24%	47%	45%	79%	34%	
Married (v. cohabiting)	97%	98%	99%	92%	100%	97%	
Only one relationship	75%	78%	69%	66%	83%	82%	
Age at first cohab/marriage	24.0	24.8	23.1	23.7	23.2	25.1	
College degree	30%	31%	28%	27%	21%	44%	
Advanced degree	27%	31%	30%	26%	25%	22%	
Income	$51,358	$58,763	$60,985	$42,443	$29,706	$55,461	
Not working	19%	16%	14%	17%	30%	26%	
Part-time	14%	4%	3%	6%	55%	13%	
Full-time	67%	80%	83%	76%	15%	61%	
Age at first child	26.5	28.5	26.3	23.3	25.9	27.9	
No children	4%	0%	0%	19%	0%	0%	
One child	11%	3%	23%	16%	10%	0%	
Two children	35%	53%	36%	33%	41%	3%	
Three or more	50%	45%	42%	32%	49%	97%	

Source: Cornell Community Study (1999-2000); 1,515 married and 17 single men and women in upstate New York.