Editorial Introduciton

Viewed from a physical scientific perspective, the data used in social science research are often "messy", and, more often than not, this is a consequence of their having been collected for purposes other than academic research. For statistical analysis, this means the "experiments" are essentially unplanned and out with the researcher's control. For visualization, which might be seen as a great help in the analysis of such data, it means that generic functions often do not produce useful graphics.

In this Case Study Humphrey Southall and Ben White start with a deceptively simple visualization problem: how can we visualize individual life histories in time and space? They illustrate two key points:

1. in fact, this is a specific case of a rather more general problem, that of visualizing longitudinal data.
2. an old and well understood graphic, the lifeline diagram developed by Swedish "Time Geographers" in the 1960's, provides a possible solution which has not been implemented this far.

Mapping the Life Course:
Visualising Migrations, Transitions & Trajectories

Humphrey Southall & Ben White

Department of Geography, Queen Mary & Westfield College,
University of London, Mile End Road, London E1 4NS, United Kingdom

Email

Web page: http://www.lifeline.qmw.ac.uk/

Abstract

This case study explores alternative approaches to the visualisation of longitudinal datasets derived primarily from life histories for large numbers of individual people. Recent years have seen a large expansion of social science research based on such data, and an associated development of tools for statistical analysis: survival analysis, logit and probit modeling, and so on. However, visualisation tools are far less well developed. The study suggests that useful ideas can be borrowed from the work of time geographers active in the late 1970s and early 1980s, and focuses particularly on lifeline diagrams, in which individual lives are represented as horizontal lines on which events are marked by point symbols and states are shown by styles or colours of line.

The Case Study begins by discussing how best to visualise an individual's history of geographical movement, via a conventional map annotated with dates or possibly by a three-dimensional plot in which the vertical axis is time. It moves on to consider geographical movement by a large group of individuals, introducing both the lifeline diagram and an extended example, a large database constructed from the membership and benefit records of a 19th century trade union, the Steam Engine Makers' Society. Via this example, the Study moves on from geographical mobility to other transitions within individual lives, particularly the movements between being in work and being in receipt of various welfare benefits. The use of similar graphics to present transitions in the lives of other entities of interest to social scientists, such as regions and nations, is briefly discussed.

The remainder of the Case Study explores the potential application of computers for visualizaton. The size and complexity of longitudinal datasets makes fully interactive graphics currently hard to achieve without using exotic hardware. However, an interactive 'lifeline viewer' could greatly assist researchers in exploring and interpreting lifelines for large numbers of individuals generated non-interactively. The essay reports on a range of contacts with researchers in the field, noting the fragmented pattern of activity and lack of specialised tools, but also the great interest and enthusiasm. It concludes by suggesting that if there are to be any new resources to develop visualisation methodologies within the social sciences, longitudinal research might benefit more than better developed fields such as Geographical Information Systems.

1. Introduction

Traditionally, quantitative social scientists have gathered data primarily through cross-sectional sample surveys carried out at single points in time, and have relied on aggregate time series primarily constructed by government statisticians to study change. Methods for both analysing and visualising such data are well-established: for cross-sectional data, the standard inferential methods, and the graphical toolkit which is included in every spreadsheet programme; for time-series data, methodologies for time-series regression which have been particularly highly developed by econometricians, and the standard time-series graphs which are again widely available.

Table 1

Statistical and Visualization methods relevant to chronological data.

Type of Data:	Statistical Methods	Visualisation Methods
Sample Surveys	Chi-Sqare, T-test, etc	Histograms, Pie Charts etc
Time Series	Time Series Regression	Time Series Graphs, etc
Panel Data/ Biographies	Logit & Probit Modeling, Survival Analysis, etc	?

However, in the last decade a great deal of research has focused on tracing individuals over time, empirically through the gathering of data through panels of individuals who are periodically re-interviewed, or asked to maintain a diary; theoretically via micro-simulation studies. There is by now a well-established body of statistical techniques for analysing such datasets — methods for event history analysis and survival analysis (see Mayer & Tuma, 1990; Courgeau & Lelièvre, 1992). However, techniques for visualising such data are far less well developed, and researchers tend to fall back on aggregating their data to create conventional time series data, or present arbitrarily chosen individual cases — in other words, we tend to have to choose between losing all the detail we have gone to such effort to assemble, or seeing just the individual trees but no wood.

Given that projects assembling life-course data are among the most expensive in British social science such as the 1958 National Child Development Study, the ONS Longitudinal Study 1971-91 and so on, it is essential that findings be not only statistically significant but communicable to a broad audience. Unless we can convey the diversity of **life-course** experience, popular notions will tend to be based on a **life-cycle** model derived from averages. One example of how misleading this can be is that in the mid-19th century life expectancy at birth could be below 50, but this was due to high infant and child mortality, not to any significant fraction of the population dying in their 40s.

This Case Study explores methods for visualising the sequence of events that make up a life, and in particular a particular form of visualisation which is seems to be generally known to its scattered users as a **lifeline diagram**. The precise origins of both this diagram and its name have not been established, but many ideas in this area can be traced back to **Time Geography**, and through that to Swedish geographers and in particular Torsten Hägerstrand. **Time geography** was, in retrospect, perhaps something of a passing fad within human geography of the late 1970s and early 1980s (eg Carlstein *et al*, 1978; Parkes & Thrift, 1980), concerned not so much with the long time spans of historical research but with daily, weekly and seasonal rhythms within human behaviour over space. Its approach was often highly conceptual, actual empirical studies were often on a relatively small scale. Its importance here lies mainly in the range of ideas it came up with for graphically portraying individual-level data involving a time dimension.

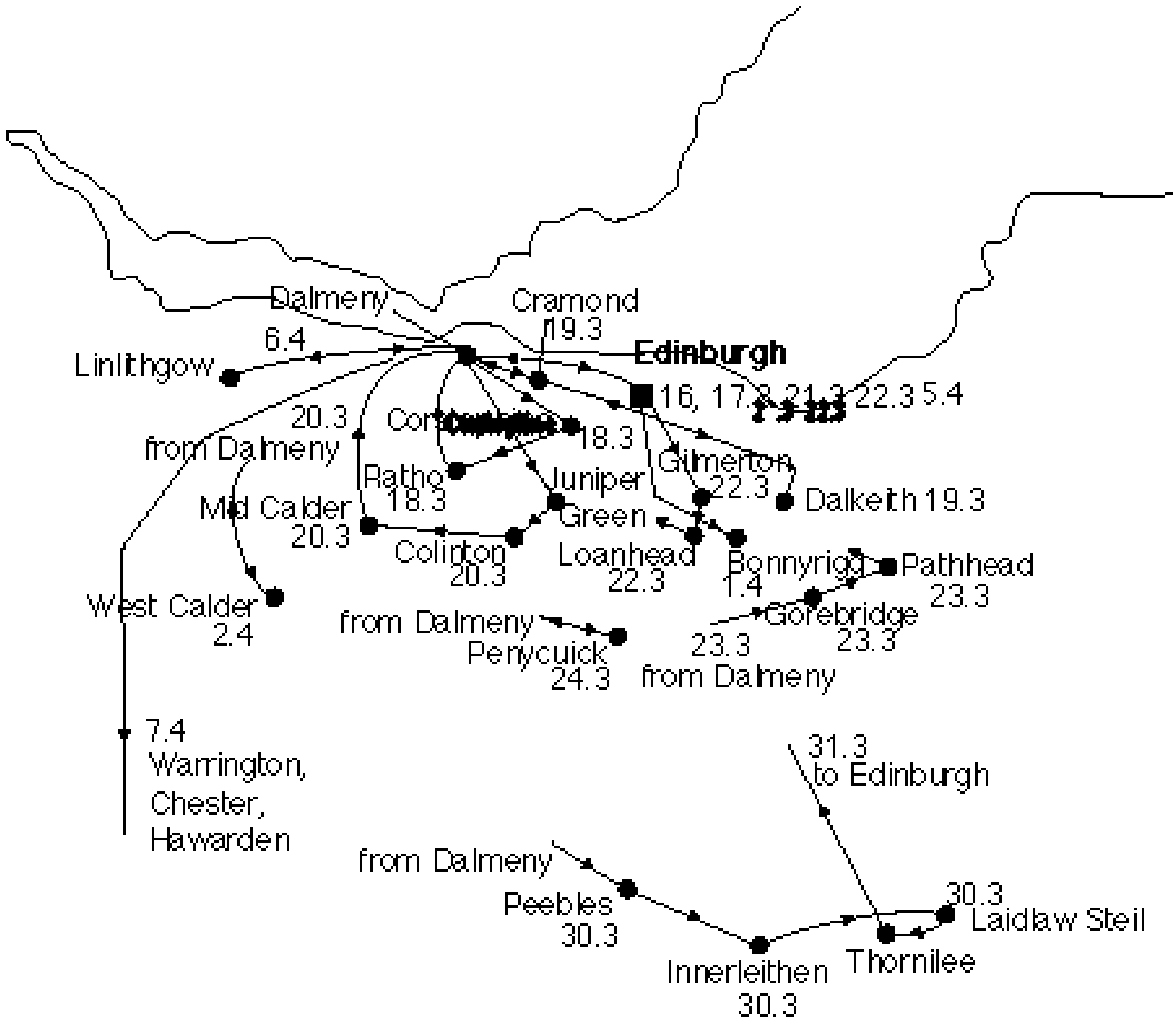
As far as we have been able to discover, the vast majority of the visualisation created by time geographers were executed entirely manually, created not by the researchers themselves but by the cartographers their departments employed at the time. The fact that geographers are so heavily involved in current visualisation research in the social sciences partly reflects the relatively generous equipment grants they enjoy, but also a longer tradition of specialist support for non-computer graphics. Even with such support, many of the visualisation methods developed by time geographers were enormously time consuming, taking days or even weeks to draw, and in practice they seem to have been rarely used.

Like the work of time geographers, the examples used in this essay are all essentially manually drawn, and this requires some defence in an initiative concerned with computer graphics. In a strict sense, most of the graphics here are computer-generated, in that they were drawn using Adobe Illustrator, but often starting from a scan of an original drawn with pen and ink. However, again as far as we can discover, no software exists to create a final lifeline diagram from raw data, and the partial solutions discussed create poor quality output. Given that the central end-product was to be a conventional essay, where the software used would be invisible; given that one of the two other non-GIS case studies within this Initiative was concerned with the creation of Lexis pencils (a form of lifeline diagram), and given that the time available meant that even if we had spent all our time developing software its functionality would have been very limited, it was decided to concentrate on exploring ideas rather than developing software.

2. Visualising Geographical Mobility

The geographical movement of a single individual over their lifetime is one of the most easily visualised of all social processes, but it is a special case of a very complex class of time-based visualisation problems involving a single actor and a two dimensional space. Even here, there is a third dimension, time, which is not always easily represented. Demographic research has tended to concentrate on relatively infrequent movements by individuals which are classed as 'migration moves', associated (hopefully) with changes of both employer and residence, but once we have the ability to trace individuals continuously much more complex patterns of movement emerge. Here we have a problem in that really detailed information about the activities of modern individuals may well be available, but they may object to it being displayed on the World-Wide Web! Let us take one well known figure from history, the late 19th century British politician William Gladstone. He left a daily diary (Matthews, 1986), and the entries for 16th-18th March 1880 look like this:

Figure 1: William Gladstone's Movements during the Second Midlothian Campaign, 1880



16. Tu. [Dalmeny House, Edinburgh] Wrote to Mr Newton – Mr ... – Scotts. Saw Herbert G. & instructed him as Private Sec. Packing & off at 9.30. Was obliged to address the people at every point (5) before Edinburgh – At **York** there were I think 6000: very quiet. At **Edinburgh** the wonderful scene of Nov. was exactly renewed. Reached **Dalmeny** 8.30 PM. Read Tracts on Mr Cowan – Pr. Consort Vol V – Tourgenieff Terres Vierges. Evening with Mr Reid & Mr Richardson on the arrangements to be made.

17. Wed. Wrote to Mr Madan – Mr Toynbee – Mr Murray – Messrs C. Scribner & Sons – Mr Adam l. & 2 Tel. Drove to **Edinburgh**. Great & most enthusiastic meeting in the Music Hall. Spoke 1hour: questions &c. followed. Worked up Probate Duty figures. Quiet evg at Dalmeny. Read England under Lord B. Began Pr. Consort Vol. V.

18. Thur. Wrote to Mr Hinde Palmer – Rev. Mr Enraght – Lord Dalrymple – Herbert John G. – Mr MacCulloch – Mr Tennant MP – T. Reeves – J. Rankin – Mr Knowles Tel. – Mr Myers WE. Meeting at **Corstorphine** in Free Kirk. Sp. 1hour. And at **Ratho** in School: hour. All signs good.

> And dear dear Herbert is out for Middlesex. God be with him! Read Lefevre on Game Laws – Life of Prince Consort.

These days were the beginning of his second Midlothian campaign, the diary provides precise associations between dates and geographical locations, and the following map reconstructs his daily movements during the whole of the campaign (Figure 1). This map may look unremarkable, but note that this was part of a General Election campaign and Gladstone was effectively running for Prime Minister (see Southall, 1996). Can we imagine Tony Blair staying inside his own constituency in the same way? In a very simple way, the map makes a point that the diary does not, unless one is intimately familiar with the names of villages around Edinburgh.

Gladstone, of course, was a well-known politician whose movements were recorded not only in his diary, subsequently published, but by newspapers. The movements of 'ordinary people', perhaps fortunately, are not systematically recorded in such detail, but there are **many** exceptions to this. One obvious exception is where researchers interview them, or ask them to keep a diary. In historical contexts, we can sometimes trace movements through autobiographies (see Southall 1996) or through the work of family historians. However, an unusual source which provides the basis for a number of examples used here are the records of an early trade union, the Steam Engine Makers' Society, which listed in their *Annual Reports* both their membership records, including transfers between branches, and details of all benefit payments to members, including a 'traveling benefit' which assisted members as they traveled around the country seeking work.

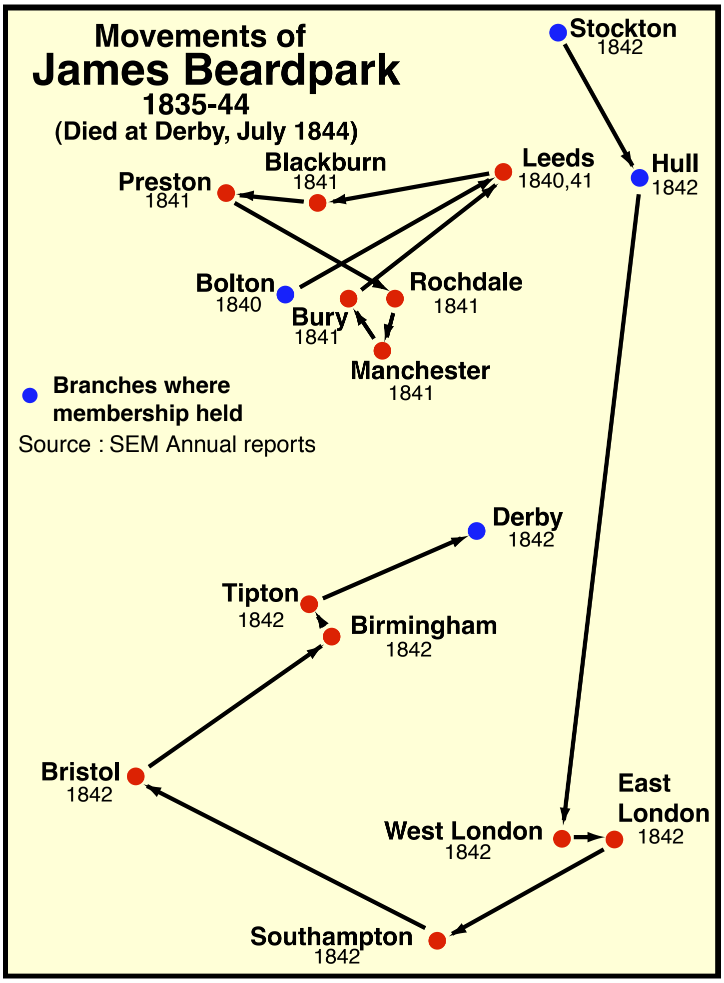
Table 1 assembles together all the information on a whole series of the union's reports about a man called James Beardpark (although his name appears in a number of slightly different ways). They trace him moving from his first being listed in the Bolton branch in 1835 to his death in Derby in 1844. The list includes various payments of sick pay (he was off work for five months in 1838 due to an accident), his wife's funeral in 1837 and his own funeral in Derby, and in particular three periods when he was tramping around the country looking for work. In April 1840 he seems to have moved from Bolton to Leeds and joined the branch. In September 1841 he left Leeds, and tramped through several Lancashire towns but then we lose track of him for a time. He reappears in Stockton but left in April 1842, going first to London (probably partly by sea) and then through Southampton, Bristol and Birmingham until he ended up in Derby. The union paid funeral benefit on his behalf two years later, and more details were obtained from his death certificate. He died at the age of 35 'by the Visitation of God of Natural Causes'.

Table 1
Information from the SEM *Annual Reports* concerning 'James Beardpark', 1835-44

Date	**Branch**	**Type**	**N^o**	**Name**	**Details**
1835/6	Bolton	Member	57	James Beardpark	
9/3/36	Bolton	Sick Pay		James Beardpark	26 days £1 19s. 0d.
1836/7	Bolton	Member	46	James Beardpark	
2/1/37	Bolton	Funeral		*Jas*. Beardpark	For wife £3 0s. 0d.
1837/8	St. Helens	Member	22	James Beardpark	Gone to Bolton
16/12/37	St. Helens	Sick Pay		James Beardpark	12 days £1 10s. 0d.
24/2/38	St. Helens	Sick Pay		James Beardpark	22 wks 4 dys £11 11s. through an accident
1837/8	Bolton	Member	123	James *Beanpark*	Has joined St.Helens
1838/9	Bolton	Member	95	James Beardpark	
27/8/38	Bolton	Sick Pay		James Beardpark	14 days £1 1s. 0d.
20/2/39	Bolton	Sick Pay		James Beardpark	26 weeks £6 16s. 6d.
3/6/39	Bolton	Sick Pay		James Beardpark	15 weeks £2 12s. 6d.
1839/40	Bolton	Member	80	James Beardpark	Drew Clearance
6/4/40	Bolton	Sick Pay		James Beardpark	21 weeks £3 13s. 6d. at 7s.6d. per week
11/4/40	Leeds	Travel		James Beardpark	From Bolton 9s. 0d.
25/8/41	Leeds	Travel		James Beardpark	From Bury 2s. 6d. Refreshments
7/9/41	Blackburn	Travel		James Beardpark	From Leeds 6s. 7d.
9/9/41	Preston	Travel		James Beardpark	From Blackburn 3s. 0d.
23/9/41	Rochdale	Travel		James Beardpark	From Preston 5s. 2d.
12/9/41	Manchester	Travel		James Beardpark	From Rochdale 3s. 4d.
18/9/41	Bury	Travel		*Joseph* Beardpark	From Manchester 8d.
1841/2	Stockton	Member	9	James Beardpark	Obtained Travel. Cert.
7/4/42	Hull	Travel		James Beardpark	From Stockton 12s. 2d.
8/4/42	W. London	Travel		James Beardpark	From Hull 8s. 0d.
9/4/42	E. London	Travel		James Beardpark	From Lambeth 5s. 4d.
11/4/42	S'hampton	Travel		James Beardpark	From London 9s. 6d.
16/4/42	Bristol	Travel		James Beardpark	From S'ton 12s. 7d.
23/4/42	Birming'm	Travel		James Beardpark	From Bristol 12s. 0d.
5/42	Tipton	Travel		*John* Beardpark	From B'ham 3s. 2d.
2/5/42	Derby	Travel		James *Bearpark*	From Tipton 8s. 4d.
1841/2	Derby	Member	33	James Beardpark	
1842/3	Derby	Member	28	James *Bearpark*	
20/8/42	Derby	Sick Pay		James *Bearpark*	5 days 8s. 9d.
1843/4	Derby	Member	15	James *Bearpark*	Is dead
22/6/44	Derby	Funeral		James *Bearpark*	£8 0s. 0d.

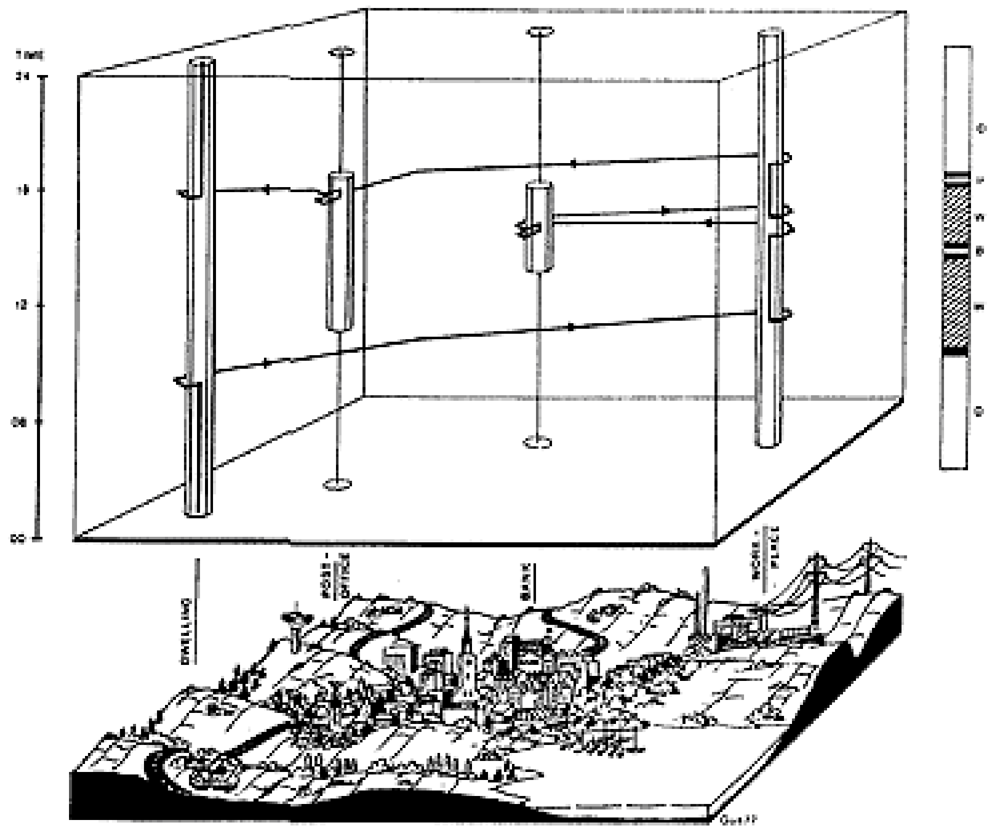
As with Gladstone, this career becomes a little clearer if we map it, as shown in figure 2. The blue dots are towns where he was listed as a member, while the red dots are towns where he received traveling benefit.

Figure 2: Movements of James Beardpark 1835-44



The problem with simple maps like those for Gladstone and Beardpark is that they do not provide an explicit time dimension, so this has to be added through annotations, and where a person visited the same location more than once the resulting map can be very confusing. Time geographers developed sophisticated graphical representations of the life course, such as the Figure 3, taken from Lenntop's chapter in Carlstein *et al*. (1978), but they have been used only rarely and at great cost in illustrators' time. The figure below represents the movements of a person over a single day. The three dimensions of geographical space have been collapsed into two in order to make room for a time axis. The movements of the individual in space and time are depicted in a continuous and indivisible path. In the example, the individual starts from the home and visits his [*sic*.] workplace, a bank, his work place and finally a post office, before returning home." The shaded bar at the right identifies periods spent traveling (in black) and at work (cross-hatched). This use of a pseudo-three dimensional representation suggests an obvious potential for computer graphics.

Figure 3: Example of an individual's path in a time-space coordinate system

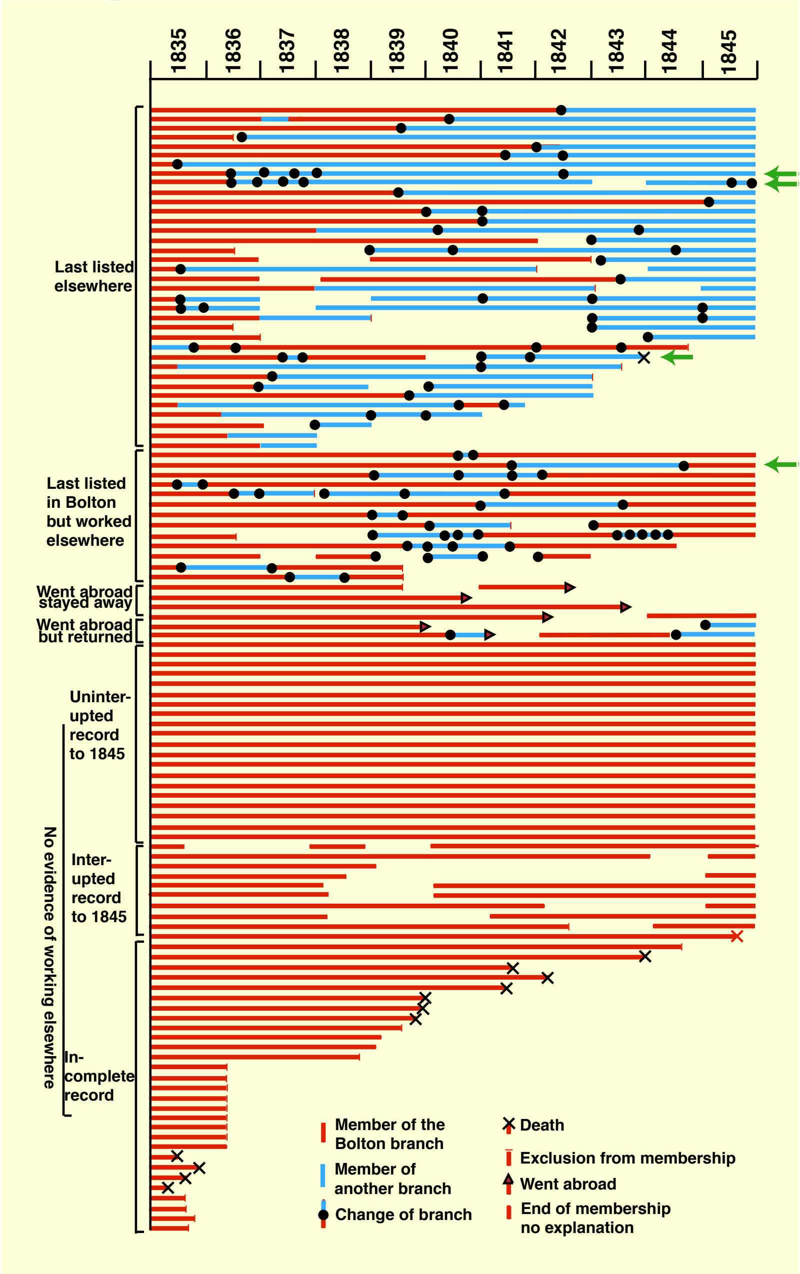


(Source: Carlstein *et al.*, 1978, p.164)

These examples concern the movement of a single individual, but how can we best present the experience of a group? If we are limited to graphics that can be reproduced on paper, we begin to run short of dimensions: if one dimension is time and another is used to distinguish individuals, how can we show where they moved? These constraints lead naturally to our first lifeline diagram, where each career is shown by a single line or bar and event are indicated by symbols or changes in fill patterns along the bar. For example, the following diagram summarises the migratory history of a large group of members of the Steam Engine Makers' Society. The Figure 4 is a simple bit-map graphic, but it forms part of a far more complex example exploring the idea of 'drill-down'.

To provide some background to this diagram, the late 1830s and early 1840s, the middle of the time-lines, were a period of severe recession (associated with Chartist unrest) and it would seem to have affected the north-west of England and towns such as Bolton particularly severely. As a result, many engineers were forced to look elsewhere. The members of this trade union were able to obtain financial assistance from other branches as they traveled around, and from this we have very precise records of their movement. Note that about half the membership worked outside Bolton at some time, and a significant fraction went abroad. Most men who moved did so more than once, and many ended up back in Bolton. In the working example, it is possible to click on some of the lifelines to obtain more information about individuals.

Figure 4: Subsequent Careers of 1835 Members of the SEM Bolton Branch



A second, very similar, example (Figure 5) of a lifeline diagram showing the geographical movements of twenty individuals is taken from Langton & Hoppe's (1990) work on rural-urban migration in mid-19th century Sweden

Figure 5: Life-time Migration histories of 20 people sampled from the *husforhorslangder* of Vadstena in 1855 (the numbers in the left-hand margin indicate occupations)

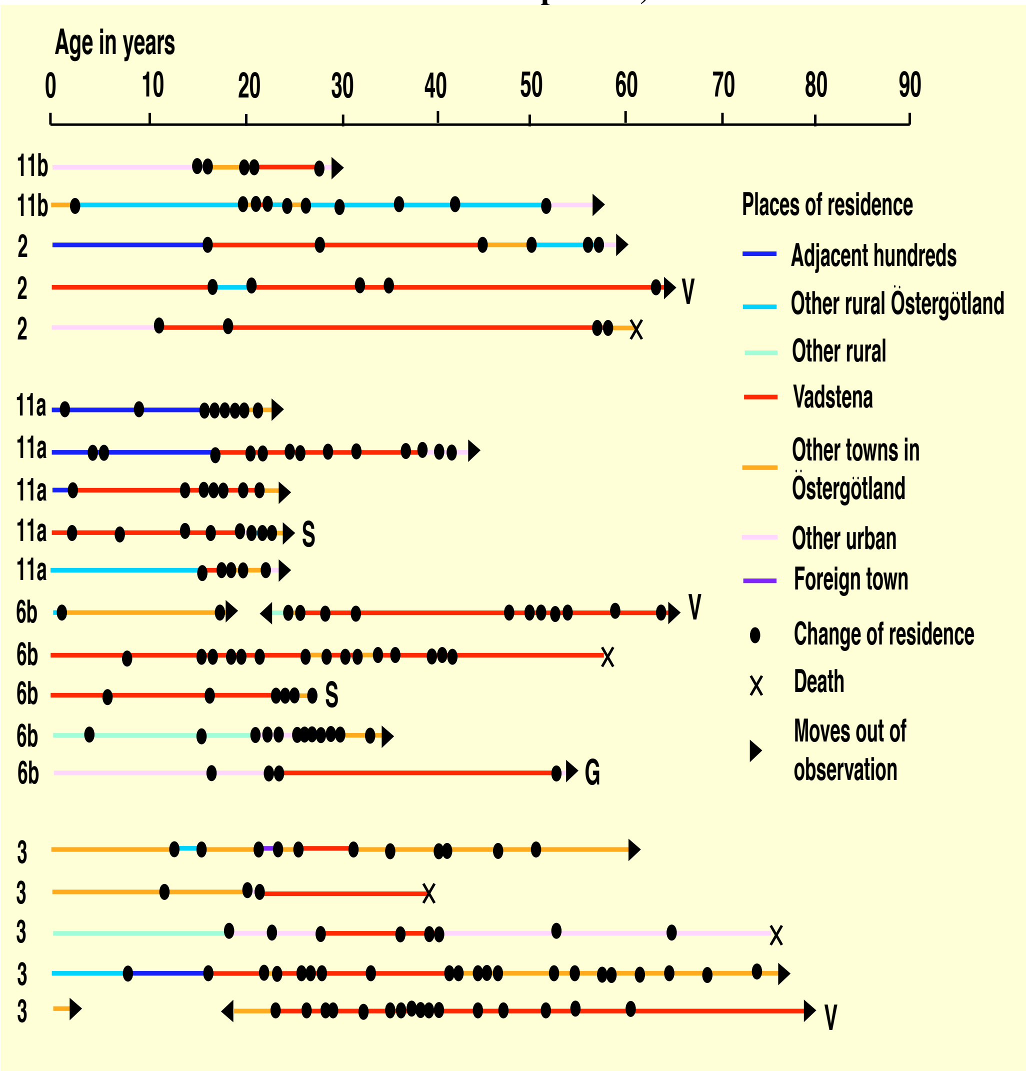
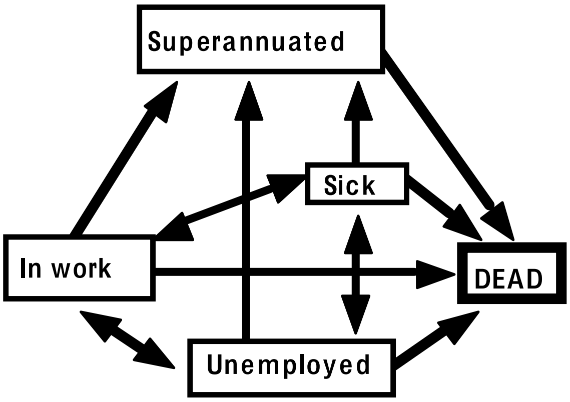


Figure 5 has been redrawn from the original figure, which covered 47 people and was limited to black and white. Rather than use colours, the original uses three weights of continuous line, three weights of dotted line and a wavy line to show the different 'places of residence'. Arguably, so many different types of line confuse the reader and it would be better to provide less detail in the main diagram and more via 'drill-down'. We have also taken the opportunity to make the point symbols more closely resemble those used in the previous example. Langton and Hoppe comment that the diagram:

> "reveals much more fully the nature of the massive residential mobility shown in the tables. It was normal for people to change their place of residence at all stages of the life-cycle. Although there was a tendency for this mobility to be at its most frenetic when people were in their twenties, everyone had moved, some two or three times, by their early teens, and everyone continued to move about until their sixties and seventies if they lived so long".

3. Visualising other Transitions in individual lives

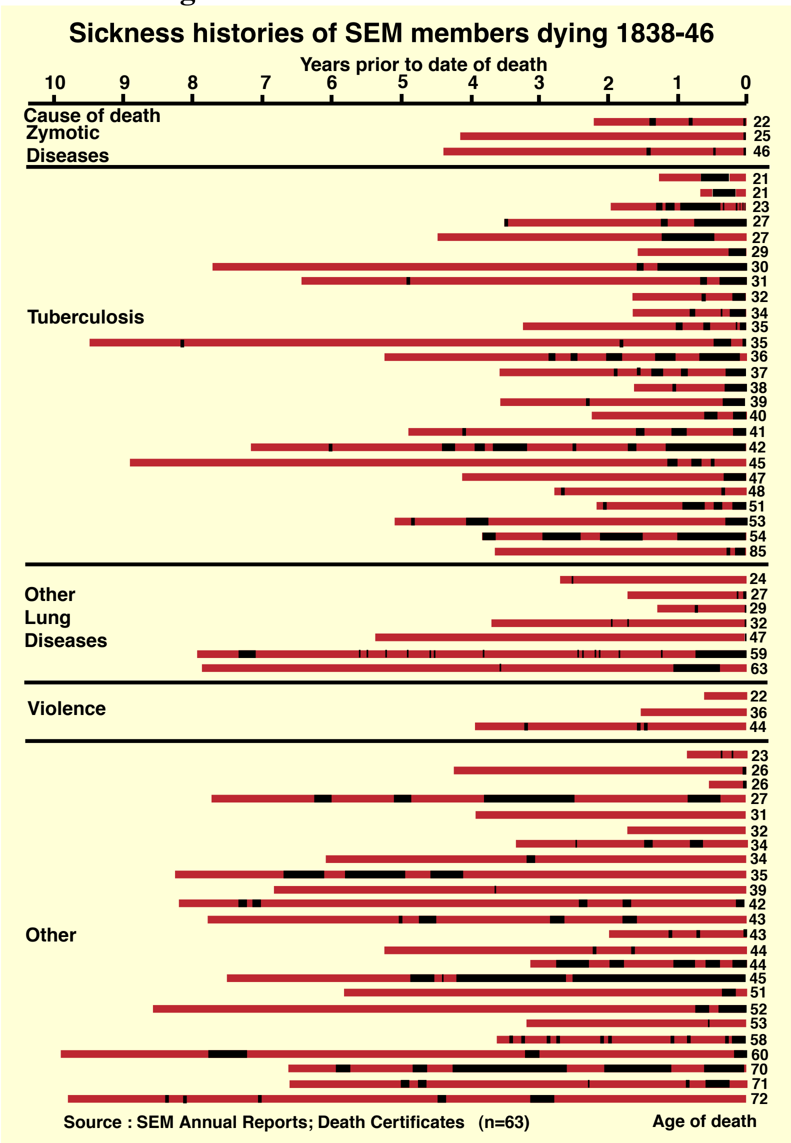
Figure 6:



The above examples all concern transitions between places or trajectories through space, but similar methods are applicable to other dimensions of the life course. The Figure 6 shows a series of possible transitions which can affect the individual worker, reducing their productivity and increasing demands on welfare systems; note that the only transition out of superannuation is death and that there are no transitions from death. The Steam Engine Makers (SEM) database, used above to study migration, records all these transitions for this particular group of C19 workers. The Figure 7 (from Southall & Garrett, 1991) shows the sickness histories for SEM members; note that time can be expressed in three ways: years from birth, indicated by the ages on the right-hand margin; years before death, shown on the horizontal axis; and, missing here, the actual calendar date. One rather obvious comment about what this figure shows is that men who died of TB tended to be sick for lengthy periods prior to their deaths, while those dying of 'violence' (mainly industrial accidents) experienced little prior sickness.

Other studies taking similar approaches are Savage (1993) on the career paths of banking workers, which he uses to explore gender differences and the emergence of the 'modern career', and Alter (1988) on the linked migratory and fertility histories of Belgian women. These examples are all 'historical', but to some extent any study of lives as a whole must span decades. The issues addressed by such life-course research are of great contemporary significance; for example, Falkingham & Hillis (1995) in ESRC-funded research explored the relationship between variations over the life-course and between lives in net benefits from state welfare systems. Differences between lives are critical, but their presentation is limited to just four samples as shown in Figure 8.

Figure 7: Sickness histories of SEM members dying 1838-46

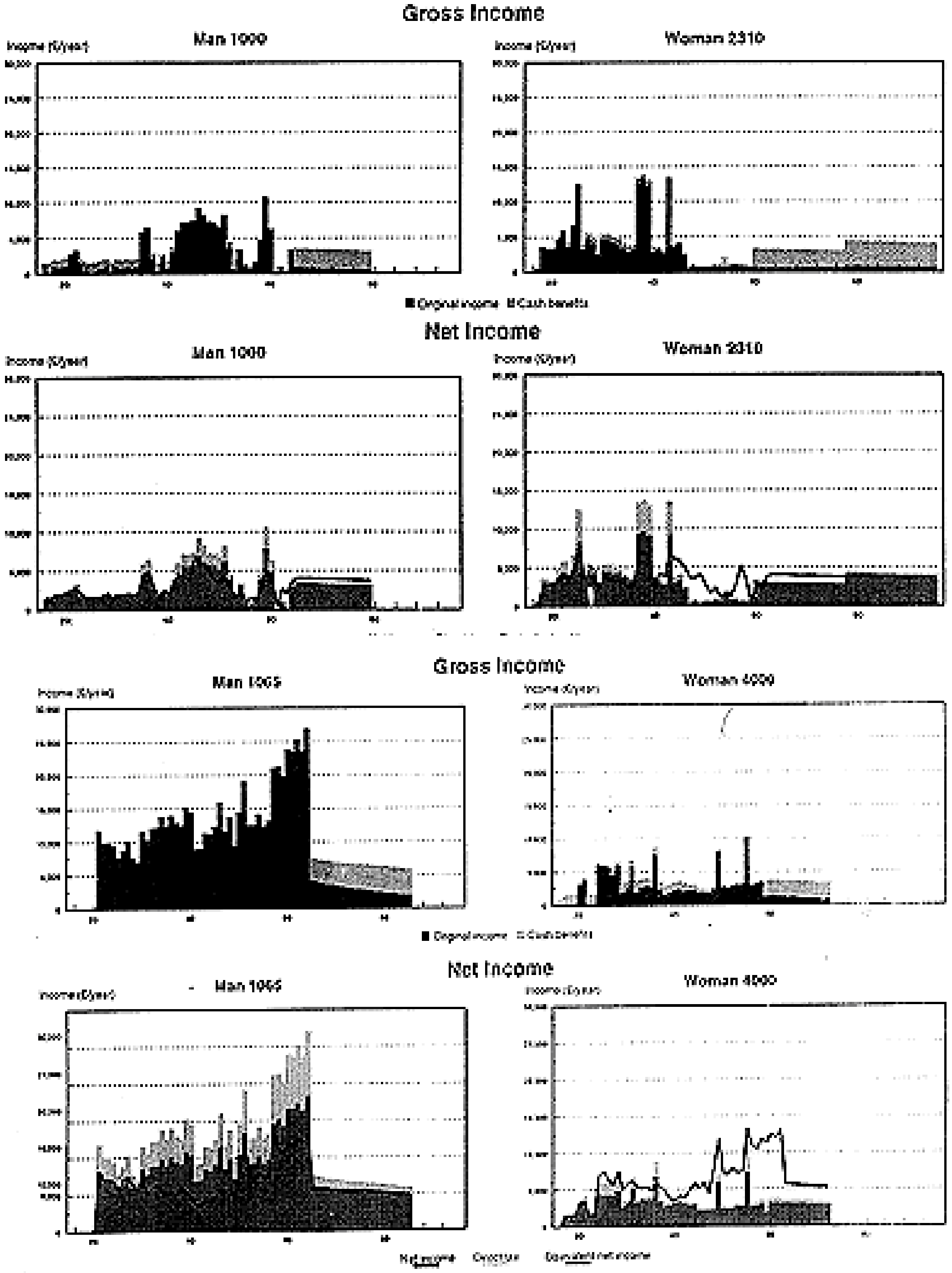


The above graphs concern **not** real people but sample output from micro-simulation models. The horizontal axis is years from birth and the vertical axis is income, either from work (in black) or from benefits (in grey). For example, man 1000 (at top right):

> "has a poor employment record. He leaves school at 16, and being unable to find regular employment, receives Supplementary Benefit (SB) from age 16 to 20. Because of his interrupted work history he only receives Unemployment Benefit in the years when he is 23 and 24 (during which entitlement runs out) and when he is 38; at all other times when he is out of the labour market he is reliant on SB. The only time he pays significant amounts of tax, reducing his net income [lower graph] below his gross income [upper graph] is during his mid-30s".

Conversely, man 1065 (at bottom right) remains in full-time education to age 20 but then receives a much higher income until retirement at age 65 (see Falkingham & Hillis, 1995 pp.77-82).

Figure 8: Gross and Net Lifetime Incomes for Two Simulated Couples



(source Falkingham & Hillis, 1995 pp.78 and 80).

4. Visualising Trajectories in the ‘lives’ of nations and regions

The examples discussed so far concern people’s lives, but the methods are relevant to any entity capable of change over two or more dimensions: firms are created, merged and de-merged, and closed down; products and political parties are 'repositioned' in the marketplace. Two final examples concern nations and regions.

Firstly, perhaps the best know 'transition diagram' in the history of the social sciences, and one of the most problematic. Nations, like people, follow a variety of paths but **Rostow's 'stages of economic growth'** (1960) allows them just one: arguably, the American Way. Even though the fallacy was demonstrated decades ago, students still arrive at university having been taught this model, arguably a testament to the power of graphics. Treatments such as Solow's (1970) may contain far more insight into the real processes of economic growth, but though it contains diagrams they provide detail not overview. Rostow's diagram shows the fundamental danger of two dimensional visualisations in this field: if one axis is time and the other distinguishes cases (just like the SEM diagrams above), there is no space within which nations can follow different paths and as with individual life-cycle models the only variation is in timing.

Figure 9: Walt Rostow's 'Stages of Economic Growth'

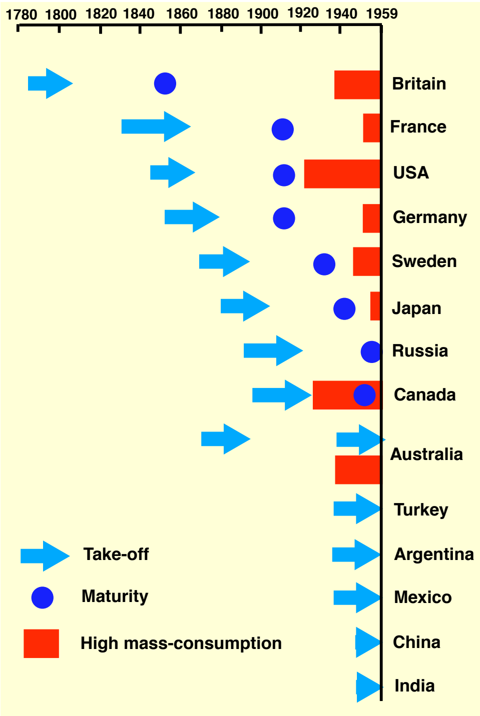
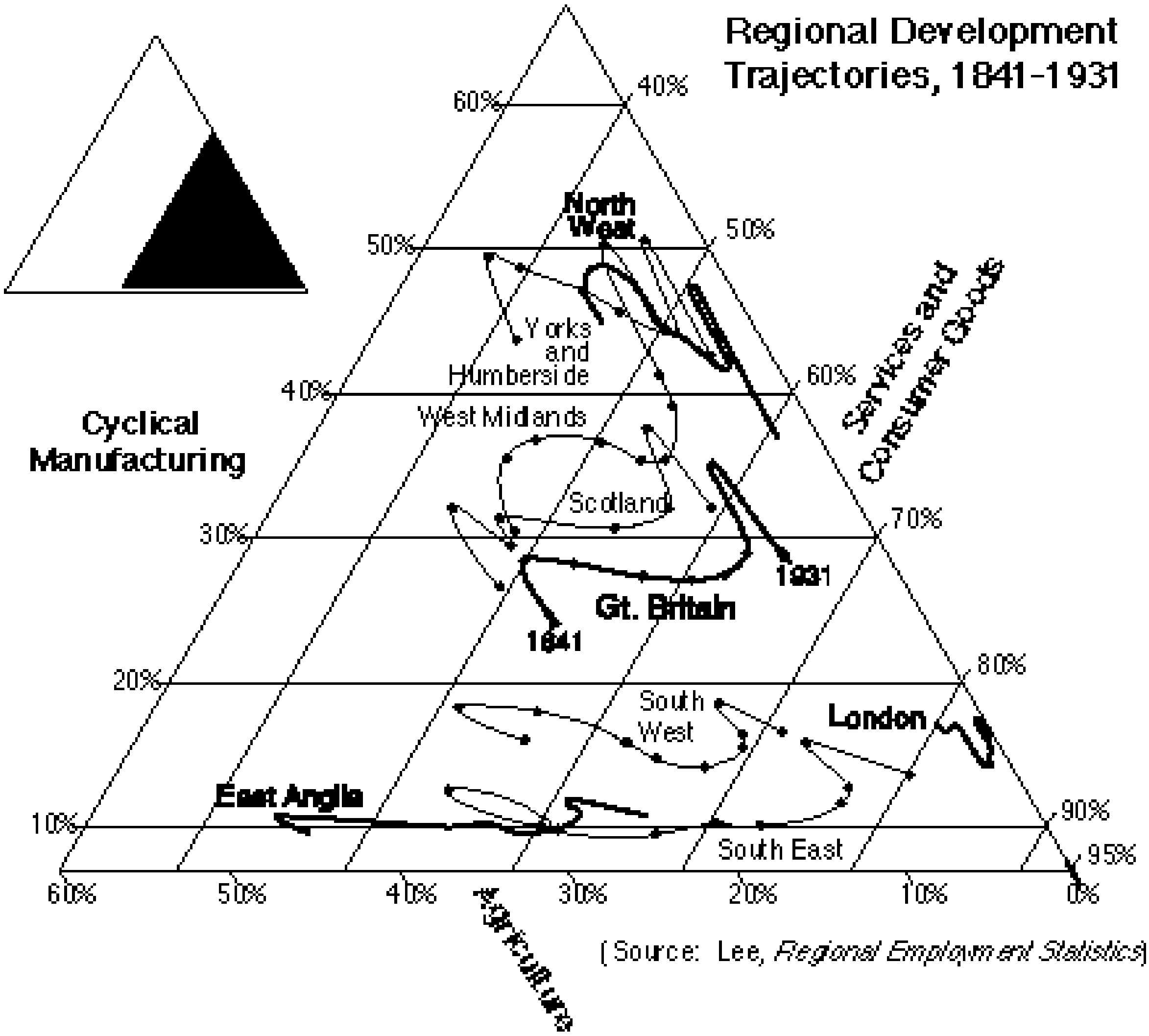


Figure 10, taken from Gilbert & Southall (1994) takes a slightly more sophisticated approach to the economic development of the regions of Britain, plotting their trajectories in terms of industrial structure between 1841 and 1931. It attempts to present five dimensions on a two dimensional surface: time, the three economic sectors and the geographical dimension which distinguishes the regions. Despite its complexity, this is still a simplification from the original statistics for 32 sectors and 52 counties. It should be apparent that each of the three regions highlighted in blue followed a markedly different path, and the national path (in red) was an average between these very different patterns.

Figure 10: Regional Development Trajectories, 1841-1931



5. Discussion: Computerising Lifelines

So far, this essay has been concerned with visualisation but not with computer graphics. The remainder of the essay explores the pros and cons of generating lifeline diagrams through software, using the Steam Engine Makers' database as the central example but also drawing on discussion with researchers working with two large scale modern longitudinal datasets:

- The Office of National Statistics' **Longitudinal Study** (LS) is based upon the census and vital event data (births, cancers, deaths) routinely collected for 1% of the population of England and Wales — approximately 500,000 individuals at any one census point. The study contains data on LS members present at the 1971, 1981 and 1991 Censuses plus information on other individuals living in the same household at each census. Academic access to the Longitudinal Study is the responsibility of the Social Statistics Research Unit at City University, and we have been in contact with both the Unit and a number of their collaborators at other universities. For further information, see: http://ssru.city.ac.uk/Ls/lshomepage.html

- The **British Household Panel Study** (BHPS) is following all members of a sample of households through repeated interviewing in a series of 'Waves', now up to wave 6. The Wave 1 panel consists of some 5,500 households and 10,300 individuals drawn from 250 different areas of Great Britain, although successive waves each lose a few more members of the panel. It is the responsibility of the ESRC Research Centre on Micro-Social Change at Essex University and more information is available from: http://www.irc.essex.ac.uk/bhps

From these contacts, a number of general points emerged:

5.1 Life-course datasets are both large and complex.

While time geographers sometimes based their work on an individual's diary, the Steam Engine Makers' database already consists of nearly 200,000 records, while adding the full run of data to 1919 would bring it close to a million records. Modern datasets such as the LS and BHPS are even larger. Further, longitudinal datasets necessarily have a complex structure, consisting of a varying number of different types of event affecting an individual at irregular intervals over time, often linked with a more conventional set of attribute data (occupation, nationality, height, weight and so on). This often involves many separate tables within a relational framework. In such a context, visualisation is as much concerned with extracting data in a usable form as with creation of a graphical image *per se*. This often involves creating a more regular data structure which records an individual's location/status at fixed intervals; for example, within the SEM database many analyses and graphics are based on a derived dataset which records whether an individual was on unemployment, sickness, superannuation or no benefit on the first day of each month. The Micro-Social Change centre describe such a dataset as a 'calendar'.

If much of the work involved in writing a lifeline diagram generator concerns creating a calendar, a task which will necessarily vary depending on the structure of the database, the development of such software may be uneconomic unless the structure of longitudinal databases can be standardised. Further, the computation involved in generating a lifeline diagram covering several thousand individuals from a relational database is such that it is hard to imagine such graphics being generated entirely interactively. If it could be done at all, it would arguably require some combination of a very large in-memory database and parallel processing.

5.2 Researchers using longitudinal datasets are attempting visualisation work, but lack specialised tools.

The Social Statistics Research Unit and the Micro-Social Change Centre contain large numbers of full time researchers but lack specialised graphics facilities. The level of interest in visualisation is therefore remarkable, but the tools being used are remarkably crude. As in other areas of the social sciences, Microsoft Excel provides a *lingua franca*, and can be persuaded to create a lifeline diagram but only once a calendar dataset (see above) has been generated, often by strictly manual methods. For example, Professor Peter Elias of the Institute for Employment Research at Warwick University has written software for graphical analysis of the National Child Development Study (sweep 5) event histories which is primarily a data management package which then interfaces to SPSS or Excel for plotting.

The most extreme example of unlikely software being pressed into service is Brendan Halpin of the Micro-Social Change centre's use of the EMACS text editor! A calendar data set is created consisting of a series of lines, one for each person, each containing a sequences of letter codes indicating the person's status at each point in time, EMACS is used to globally append escape codes changing the background colour depending on the letter that appears. The result is a coloured lifeline diagram similar to those in this essay — and so long as you know the escape codes EMACS is much faster than Excel.

5.3 Lifeline diagrams on paper are generally overloaded; computers could help.

The published lifeline diagrams reproduced in this essay tend to be overloaded, containing a jumble of information in an attempt to provide a full record. For example, figure 5 attempts to show many different places of residence by styles of line (and the original does not even have colour), while the numbers in the left hand margin indicate occupations; figure 6 sorts individuals by cause of death, shown in the left-hand margin, and then by age, shown in the right; different lifeline diagrams organise time differently, figure 4 uses calendar years, figure 5 years from birth and figure 6 years prior to death.

While the interactive generation of lifeline diagrams covering large numbers of individuals may be impractical, an interactive tool for manipulating such diagrams, and associated information about individual characteristics, seems quite feasible. It should be able to move individual lifelines around within a viewer in two ways: Firstly, it should be able to sort the lifelines vertically by various criteria: by occupation and then by age, or by occupation and cause of death; NB this is not too hard to achieve within Excel. Secondly, it should be able to vary the basis for the time axis, and move lifelines horizontally to fit; for example, if figure 6 could be rearranged to use calendar years, it might be possible to identify epidemics.

5.4 Visualisation work may be designed to convince the 'consumers' of research of some conclusion, or to enable the researcher better to know their dataset and form hypotheses; the two roles point to different types of tool.

Even with a few thousand lives in the Steam Engine Makers' database, any graphical presentation including all the individuals or even a substantial subset will overwhelm the consumer; this case study has tried to present comprehensible examples, but it would have been easy to include many 'spaghetti' diagrams composed of endless intersecting and superimposed lifelines. In practice, users must be presented with a summary in which individual lives are aggregated in some way; arguably, this is best done using relatively conventional statistical methods, although visualisation methods may be relevant to presenting the resulting parameter estimates. One obvious example is the traditional hazard curve, expressing changes in the probability of some transition over time; another is demographic charts which summarise individual experience by comparing birth cohorts (see, for example, McKnight (forthcoming) and Anderson (1990)).

If this argument is accepted, the main users of lifeline diagrams and similar apparatus should be the researchers themselves. Here the need is to reveal complexity and, for example, the influence of exceptional cases, not to conceal them. One possibility, further explored on our web site, is a drill-down system. For example, figure 4 concerns just the Bolton members of the SEM, and might be revealed when a researcher clicked on the symbol marking Bolton on a map showing mobility rates in the different towns covered by the Bolton; the data for such a map exists and, for example, men first recorded in Bolton were markedly more mobile than Londoners. Finding that the average rate of mobility of Bolton members reflected a polarisation between men who never left and those who moved repeatedly, the researcher might wish to examine the detailed history of some of the latter. On our web site, clicking on the line within figure 4 marked by the third green arrow down takes users to the life history of James Beardpark, and to figure 2; a further click on Derby within figure 2 brings up a scanned image of Beardpark's death certificate. These samples were prepared manually, but a programme already exists which can create a textual life history for a specified member by repeatedly querying the database, and we aim to make this accessible over the web. The cost of such a system may be hard to justify for the SEM database, but would be a relatively small part of the cost of providing researchers with resources such as the British Household Panel Study.

5.5 The best-resourced developments may be taking place in the private sector.

Several people we met suggested that the high cost of creating visualisation systems which had to be tailored to work with specific databases meant that the most interesting work was likely to be going on in commercial organisations. For example, supermarket loyalty cards, which identify individual shoppers each time they pass through a store's checkouts, are leading to the assembly of vast longitudinal datasets, covering each item purchased by each shopper, with locations and dates, and linked to data on individual characteristics gathered when the customer joined the scheme and, via home addresses, to socio-economic profiles derived from the census and similar sources. What tools are being used to exploit this data?

Academic links to commercial research are much weaker here than in, say, molecular modeling. However, one interesting example was provided by the Institut für Verkehrswesen (Institute for Transport Studies) at the University of Karlsruhe, Germany. Their work on the German Mobility Panel involves specially written software for generating essentially lifeline diagrams from very short term data on individual movements, similar to that conceptualised in figure 3 (see Chlond and Lipps, 1997).

6. Conclusions

This essay was partly a return to some very old haunts; for example, the original version of figure 4 was hand-drawn in the early 1980s. However, it became something of a voyage of exploration, taking us away from our current work in Geographical Information Systems where talk of visualisation is routine into fields where individual researchers are working, often in a fairly isolated way and with extemporised software tools, to gain some insight into their data via graphics. However, the work of these researchers is arguably far more central to the social sciences than most current research in GIS, and this essay is partly based on contacts with two of the largest dedicated centres for social science research in Britain.

The time-scale of the visualisation initiative means this is inevitably a superficial tour which can make no claim to be a systematic survey of current activity in the visualisation of longitudinal data. Even so, the range of activity, and the enthusiasm of many individual researchers was remarkable, despite the lack of either graphical training or specialised software. If resources are available to further develop social science visualisation in the UK, a strong case can be made for emphasising tools for use with longitudinal data, even allowing for the large problems posed by the lack of standardisation of the underlying databases.

7. Acknowledgments

An early version of this essay was made available on the web in late December 1997, and requests for comments and further suggestions were sent out on a number of mailing lists. We are grateful to everyone who responded, especially those who sent samples of their work. We are particularly grateful to the Social Statistics Research Unit at City University and the ESRC Research Centre on Micro-Social Change at Essex University who very hurriedly organised seminars/workshops, on February 3rd and 9th respectively. We have tried to take on board the main ideas arising out of this correspondence and those meetings, but time pressure has meant that many ideas arising out of the project need to be further explored; we very much hope that later versions of this essay will be able to do this.

The Steam Engine Makers' database has its origins in research carried out by Humphrey Southall at the University of Cambridge in 1977-9 while on an SSRC studentship, and he is grateful to John Dawson, Yeshe Zangmo and the Literary and Linguistic Computing Centre for data entry assistance. The database was greatly extended in 1993-5 with funding from the Wellcome Trust and the ESRC; Steve Sargent of QMW Computing Services made a large contribution to the development of the database and associated linkage tools, and Maureen Watters was responsible for data entry and record linkage.

Many of the illustrations in this essay were originally drawn by Ed Oliver, cartographer in the QMW Geography department.

8. References

Alter, G. (1988) *Family and the Female Life Course: The women of Verviers, Belgium, 1849-1880* (Madison: U. of Wisconsin Press).

Anderson M. (1990) 'The social implications of demographic change', ch.1 in F.M.L.Thompson (ed.) *The Cambridge Social History of Britain 1750-1950 — Vol.2: People and their environment* (Cambridge: CUP), pp.1-70.

Carlstein, T. *et al* (1978), *Human Activity and Time Geography* (London: Edward Arnold).

Chlond B. and Lipps O (1997) 'Additional information out of the context: the usefulness of panel data for the correction and interpretation of different data sources' in: University of Leeds, *Papers from the Fourth International Conference on Survey Methods in Transport* (Leeds).

Courgeau, D., and Lelièvre, É., (1992) *Event History Analysis in Demography* (Oxford: OUP).

Gilbert, D.M. and Southall, H.R., (1994) 'Les multiples dimensions de la misere: les manifestations de la pauvrette d'origine economique en Grande-Bretagne, a la fin de la periode Victoriene', in M. Mansfield, R.Salais and N.Whiteside (eds.) *Aux sources du chomage* (Paris: Editions Berlin), pp. 213-50.

Falkingham, J., and Hillis, J. (1995) *The Dynamics of Welfare: The Welfare State and the Life Cycle* (Hemel Hempstead: Prentice Hall).

Langton, J., and Hoppe, G., (1990) 'Urbanization, social structure and population circulation in pre-industrial times: flows of people through Vadstena (Sweden) in the mid-nineteenth century' in Corfield, P.J., & Keene, D. (eds.) *Work in Towns 850-1850* (Leicester), pp.138-63.

McKnight, A. (forthcoming) 'Low Wage Mobility in a Life Perspective', ch.4 in Asplund, Sloane and Theodossiou (eds.) *Low Pay and Earnings Mobility in Europe* (Farnborough: Edward Elgar).

Matthews, H.C.G. (ed.) (1986*)* ***The Gladstone Diaries with Cabinet Minutes and Prime-Ministerial Correspondence***, Vol. IX, *January 1875-December 1880* (Oxford: Clarendon Press).

Mayer, K.U., and Tuma, N.B., (1990) *Event History Analysis in Life Course Research* (Madison: U. of Wisconsin Press).

Parkes, D.N., & Thrift, N.J. (1980) *Times, Spaces, and Places: A chronogeographic perspective* (Chichester).

Rostow, W.W. (1960), *The Stages of Economic Growth* (Cambridge: CUP).

Savage, M. (1993) 'Career mobility in banking', in Miles, A., & Vincent, D. (eds.) ***Building European Society*** (Manchester UP).

Solow, R.M. (1970), *Growth Theory: An Exposition* (Oxford: OUP).

Southall, H.R. (1991a) 'The Tramping Artisan Revisits: Labour mobility and economic distress in early Victorian England', *Economic History Review*, II, Vol.44, pp.272-96.

Southall, H.R. (1991b) 'Mobility, the Artisan Community, and Popular Politics in early nineteenth century England', in G.Kearns & C.W.Withers (eds.), *Urbanising Britain: class and community in the nineteenth century* (Cambridge UP, 1991), pp. 103-30.

Southall, H.R. (1996) 'Agitate! Agitate! Organise!: Elements of an historical geography of political life', *Transactions of the Institute of British Geographers*, Vol. 21, 177-93.

Southall, H.R., & Garrett, E. (1991) 'Morbidity and Mortality among mid-Nineteenth Century Artisans', *Social History of Medicine*, vol.4, pp.231-52.