## Editorial Introduction

A very common data type in social research is the categorical in which individual cases are coded by some attribute (gender, place of residence, work status, etc.). Sometimes such a variable might be an attempt to locate the individual on some underlying scale that can be assumed continuous such as "social class", but more often than not the categories are strictly nominal and many of arithmetic operations do no apply to them.

Methods for the statistical analysis of such data have been developed for many years and are nowadays widely used, but methods for their visualization available in most packages are primitive. Typically these may consist of a bar chart or "pie" diagram (often misused) and not much else, and the difficulty of visualization is compounded if the time change in such categorical variables also needs to be visualized.

Brian Francis and John Pritchard's ***Case Study*** presents the Lexis pencil as a way of visualizing a moderate number of individual time changes in a number of categorical variables. The technique and its development is of interest for its own sake, but also by the way it illustrates three more general points of interest:

1. In common with several other visualization techniques developed in this series of case studies, it builds on a well known, tried and tested graphical device, the Lexis diagram. It may well be that this types of background is an almost necessary condition for a "new" graphic to be accepted by a user community;
2. Although the graphic in static form is useful, its analytical usefulness is increased immeasurably by interaction in zooming and animation. Again, this seems to be a feature of all the successful graphic devices in these case studies;
3. Although the original intention was to build a syustem in an existing visualization system (AVS) in practice time and experience has shown that this proved very awkward and has led to the development of a stand alone program. Regrettably, this also seems to be a feature of almost all the other case studies, and it has the undesirable consequence that, in order for visualization methods appropriate to social science applications to be developed, considerable programming skills seem to be necessary. The implications of this fact for the further development of visualization in the social sciences should be obvious.

# Visualisation of historical events using Lexis pencils

**Brian Francis and John Pritchard**
***Centre for Applied Statistics***
***Fylde College,***
***Lancaster University***
***Lancaster LA1 4YF***
***United Kingdom***

E-mail: B.Francis@Lancaster.ac.uk

Web Page: http://www.cas.lancs.ac.uk/alcd/visual

## Abstract

This Case Study reports on the visualisation methods and tools developed under the ESRC Analysis of Large and Complex Datasets initiative for exploring such datasets. Longitudinal data collected over time typically consists of a large number of variables collected over a set of individuals - many of these variables will vary over time and will commonly be categorical in nature. Such social surveys will often have a start event, such as the date of marriage or the date of first criminal conviction, as well as temporal information on the individual such as age or date of birth, and this can be used in lining up or registering the pictorial representations of individuals.

The approach uses a pencil metaphor to represent an individual. Each face of a pencil represents a time-varying variable such as housing tenure, employment status or family composition - changes in state over time in these variables are represented by changes in colour along the faces of the pencil. Groups of individuals can be examined by positioning the pencils side by side ranked by some order such as number of criminal convictions or case number. However, where possible, it is often better to view the pencils in three-dimensional space by positioning the start of each pencil on a base plane defined by temporal variables such as age, calendar time or date of marriage. This has the effect of grouping similar individuals together. The method generalises the ideas of Lexis, a nineteenth century German demographer, and such plots are therefore known as Lexis pencil plots.

Special visualisation software is necessary to construct such images, and to provide the rotation, panning and zooming which the investigator needs fully to examine the dataset.
This case study used the AVS system, and a special module in AVS was written. This software tool allows the user to view any event history dataset, retaining control of the colour mapping, the width of the pencil faces, and the assignment of variables to faces. It allows subsets of cases to be viewed, cases to be identified, and selections of cases to be made partially transparent. Another necessary feature which the tool provides is the ability to rotate the set of pencils each around their its own axis while keeping the viewpoint fixed, to allow different sides of the pencils to face the viewer. The software provides hard copy output of any user view and can also write VRML format files to allow three-dimensional visuals to be disseminated.

Two short examples are presented. The first examines a criminological data set of sexual offenders and the second gives insight into a study of the employment of married couples. These studies show that Lexis pencils provide a useful way of viewing event history data and that practical insights can be obtained into complex datasets.

## 1. Event history data - the nature of the problem

In the social sciences, in considering historical information on individuals, we are primarily interested in **events** and **states**, the associations between these and changes in patterns over time. As an example, we can consider a criminal career of an offender, collected since the individual reached the age of criminal responsibility. The career will consist of a sequence of offences (events) at various dates throughout the offender's life, together with information on dates of court appearance events, the sentence passed for each offence, movement into and out of prison (states), and so on. Other information may also be collected relating to the personal life of the offender. For example, information on his marital history state (single, cohabiting, married, divorced etc.), family history (number of children in the household) and work history (unemployment or employment state, salary) might also be collected over time. This is an **event history,** and collections of individuals form an **event** history dataset. Although most examples arise naturally in the social sciences, there are many examples from medicine (examination of medical histories of a group of patients, looking at drug and surgical treatment, recurrence of disease etc.), from management sciences (time-management studies) and psychology (observations of babies after stimulus).

## 2. The need for new visualisation methods for event histories

Methods for the statistical analysis of event history data have been extended dramatically in the last twenty years. In the example above, models could be developed for the probability of reoffending or reconviction as a function of the age of the offender, current family and work history, and previous criminal history. This in general will lead to survival model methodology, and if multiple reoffending within an individual is taken into account, to survival frailty models where there is an additional unknown factor - the frailty - measuring the propensity of the individual to reoffend.

However, there has been little development of graphical methods to visualise an event history dataset before analysis. A recent book on event history analysis (Blossfeld *et al*., 1989) devotes less than 5 pages to the graphical presentation of event history data. A state-of-the-art book on statistical graphics also presents only one graphic on event data (Cleveland, 1994). This is partly because of the complexity of such datasets; it is easy to be overwhelmed with the number of variables and different dimensions of a typical study. This work intends to rectify this imbalance by exploring the potential of modern interactive scientific visualisation systems for the initial graphical examination of complex event history data. By this, we mean that we want to provide graphical displays which can display the full complexity of a dataset if this is what is required.

## 3. The development of visualisation tools

We begin by considering the types of variable which might exist in an event history dataset. First, there are variables as proxies for time, such as age of the individual, and calendar date. Second, there will be usually a large number of time-varying variables, representing changes of state or value over time. These may be continuous (such as earnings/week), ordinal (such as educational level reached) or categorical (such as type of crime committed). Third, there are time-constant variables, which do not change over time, such as the gender or ethnic group of the individual. Lastly, there are pure events, either internal to the history of the individual (such as gaining a driving licence) or common to all individuals in the study (such as a change in government).

How do we best represent an individual history? When examining the sleeping and feeding patterns of a recently born child, Cleveland (1994) considers a simple block diagram.

However, Cleveland’s idea cannot easily be extended to examine multiple histories. Barry, Walby and Francis (1989) devised a circular 'tulip' diagram, consisting of a number of concentric rings, each representing a different variable, with colour, shading and width of the rings being used to represent changes in values over time. However, such a diagram places undue visual emphasis on the variables contributing to the outermost rings, and this is not a desirable characteristic. Examples of each of these displays can be found on our web site (http://www.cas.lancs.ac.uk/alcd/visual/)

We consider instead the metaphor of a pencil, with each ot its faces representing a different variable, and its length representing the length of the event history. Calendar time runs along the length of the pencil. Changes in each of the face variables can be represented by size, value, texture, colour, orientation or shape. Categorical variables can be represented by a set of colours, textures or patterns.

### 3.1 Case Study 1 - Employment in Kirkcaldy

As an example, Figure 1 shows a perspective view of a typical pencil representing an employment event history for a married couple taken from a retrospective survey of 188 living in Kirkcaldy, Scotland in 1985 (for further details of the dataset see Francis and Fuller, 1996). The event history for the couple begins at the date of marriage at the left hand side of the diagram, and continues until the survey date.

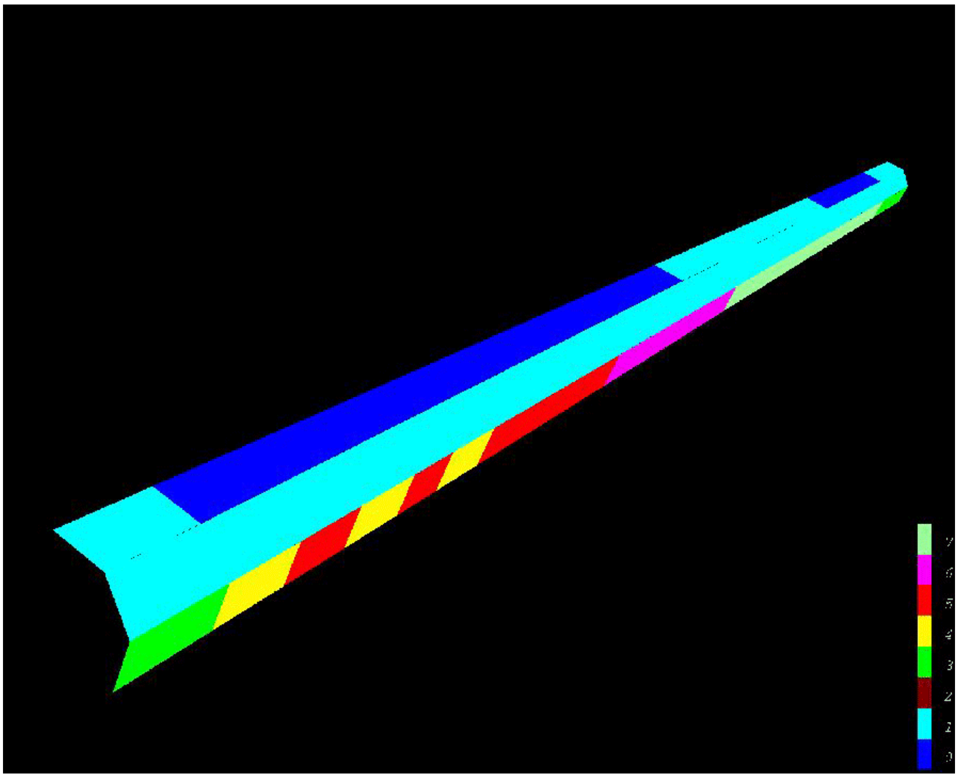

*Figure 1: A perspective view of a pencil representation of the life history of a married couple. Time runs from left to right, starting at date of marriage and finishing at the survey date. Each face of the pencil represents a different variable. The top face represents the employment history of the wife, the middle face that of the husband, and the bottom face the age of the youngest child in the household. Explanation of the colours can be found in the text.*

The three faces, taken in order, and proceeding clockwise from the top of the pencil, are female employment status, male employment status, and the age of the youngest child in the household. Employment status for both husband and wife is coded 0 (dark blue) for not working and 1(mid-blue) for working. The age of the youngest child in the household is coded 3 (green) for no children, 4 (yellow) for a child aged under 1, 5 (red) from age 1 to under 5, 6 (magenta) from age 5 to under 11 and 7 (light green) from age 11 to under 16. In this history, we can see that the husband never worked throughout the survey period. However, the wife worked up until the birth of her first child, then stopped work until her youngest child was aged 10, when she returned to work, becoming unemployed for a while in the 1981 recession.

Other analysts might want to examine other variables over time, such as migration, housing tenure, educational level of the husband and wife, and so on. The faces can be reassigned to new variables or extra faces can be added to the pencil display.

## 4. Comparing pencils - the Lexis diagram

Most studies consist of event history information from more than one individual, and we thus need methods to allow us to display more than one individual in the same graphic. The simplest method of comparing pencils is to rank them by case order, and display them side by side. This is similar in concept to a set of pencils in a pencil box, but uses no further information in the dataset. A straightforward extension to this idea is to align the pencils according to the age of the individual or according to calendar year; this then allows comparisons to be made more easily between pencils.

However, it is easily seen that these two simple displays are straightforward applications of the Lexis diagram (Lexis, 1875), the modern form of which is described in Pressat (1961). The Lexis diagram is used extensively in demography and in survival analysis and has many useful statistical properties described in detail by Keiding (1990). A typical application would look at the survival experience of a group of patients entering a clinical trial., The *x*-axis represents the calendar date, and the *y*-axis represents the time spent in the study. Each patient is represented by a solid line sloping at 45 degrees; the line slopes as both *y* and *x* increase with time. The line is anchored on the *x*-axis according to the date that the patient entered the trial. Figure 2 shows a typical Lexis diagram, showing eight individuals, with varying survival times. One of the individuals enters the study at time *T* days, and stays in the study until time *T+A* days. The time spent in the study (the y-value) is thus *A* days.

The Lexis diagram can be modified in various ways. First, information on events such as death can be added by placing an appropriate symbol at the end of the line. Second, changes of state can be introduced by using different colours to represent each. Finally, the definition of the *x*-axis can be modified. If we change the *x*-axis to represent 'date of entry to the study' rather than 'calendar date', then the solid lines will then be vertical and will no longer slope at 45 degrees, as the patient proceeds in the study, *y* increases but *x* now stays constant. The *x*-axis can also represent other temporal variables (such as age of entry into the study) or any other variable (typically these will be ranks of temporal variables such as 'rank order of age of entry into the study), but might simply be a case ordering. Of course, if rank orderings are used to define the x-axis, the special statistical properties of the Lexis diagram are lost.

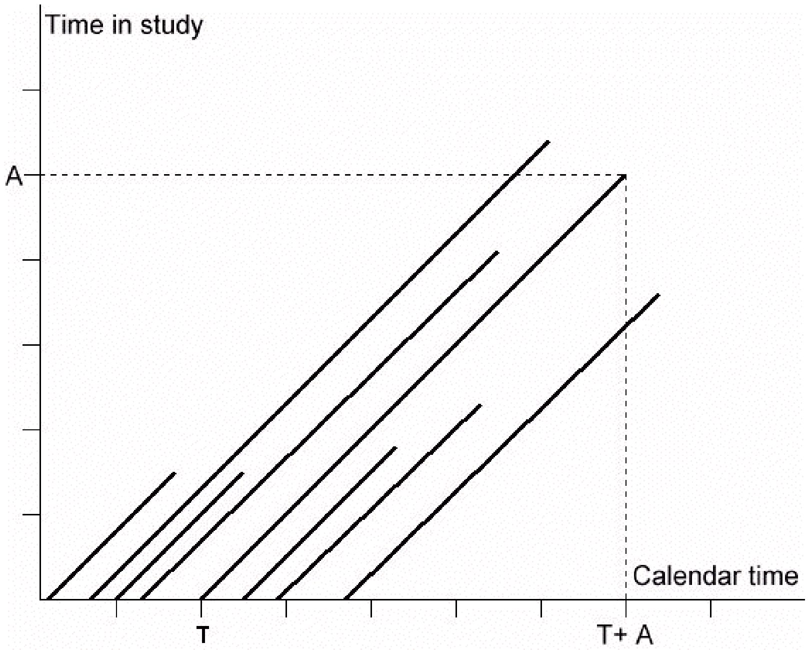


*Figure 2: A typical Lexis diagram. Each individual is represented by a 45 degree line.*

We can see that the two simple ideas for comparing event history pencils discussed above are therefore simply special forms of the Lexis diagram, with differing definitions of the *x* and *y*-axes, and with the lines replaced by pencils. In the first, the *x*-axis is the 'case order' or *index* of the individual and the *y*-axis is 'time since start event'. The second graphic redefines the *y*-axis to be 'age at start event' or 'calendar time at start event', again keeping the *x*-axis as the index of the individual.

Returning to the original concept of a Lexis diagram, and replacing the Lexis lines by pencils we can see that a 'Lexis pencil' graphic would use a temporal variable such as 'age' or 'calendar date' along the *x*-axis, and use 'time in study' or 'calendar time' on the *y*-axis. We define this to be a **two dimensional (2-D) Lexis pencil display** - the dimensions refer to the number of axes.

### 4.1 Case study 2 - A sample of bigamists

As an example, we consider the 42 bigamy offenders in England and Wales in 1973, and examine their criminal history (which was obtained from the UK Home Office Offenders Index) over a thirty-two year period from 1963 until the end of 1994. The 42 individuals consist of 39 men and 3 women, with ages ranging from 20 years to 53 years at the time of the 1973 conviction. We examine this dataset using a 2-D Lexis pencil display. The *x*-axis is defined to be the rank order of the age of the individual at the 1973 conviction, and the *y*-axis is defined to be the time since the 1973 conviction. We display a *single* pencil face which represents the principal offence at conviction. Whenever an individual is convicted, a band of colour represents the type of conviction, and the remaining time the pencil face is grey. The criminal histories are displayed from their first conviction to their last conviction within the 32 year period. Code 1 represents violence offences and is assigned the colour medium-blue, code 2 sexual offences (brown), code 3 burglary (mid-green), code 4 robbery (yellow), code 5 theft (red) and code 6 fraud and deception (magenta).

Figure 3 shows the resulting display. At $y = 0$, all individuals have an offence displayed - this is their target bigamy conviction. Mostly the principal offence displayed is a sexual offence(code 2) although for ten cases other concurrent convictions, mainly of violence(code 1) and fraud (code 6) were judged more serious than the target bigamy conviction. What stands out from this display is that there are very few cases with other sexual convictions - only 2 individuals were convicted of another sexual offence (and in neither case was this another bigamy offence). For 16 individuals, the 1973 bigamy conviction was their only conviction. However, surprisingly, of the 26 individuals with other criminal convictions, 11 of these had principal convictions for fraud and forgery (code 6) - 27% of the whole sample. Finally, there seems to be little effect of age, with no obvious change in offence specialisation when tracking from the left of the figure to the right (that is from the youngest to the oldest bigamy offender). The initial analysis of this data raises questions as to whether bigamy should be considered as a sexual offence (as the UK Home Office currently classify it) or whether it is better classified as a fraud offence (Soothill *et al.*, 1997).

The 2-D Lexis pencil displays work well when the number of individuals is small, but with larger numbers of individuals, overlap of the pencils can easily occur. How can the display be improved?

## 5. Further development - Extending the Lexis display into three dimensions

First, we note that the 2-D Lexis pencil display is rather a strange concept, with 3-D pencil objects being plotted in a 2-D co-ordinate system. An obvious extension is therefore to use a 3-D co-ordinate system to position the pencils in space, with the y-axis defined as before and representing time spent in the study, but with a base plane defined by the *x* and *z*-axes rather than a base x-axis. We define this to be a 3-D Lexis pencil display. This extension into three dimensions corresponds to the characteristics of many datasets. Often event history studies have more than one way of representing time; typical temporal variables will include the age of the individual and calendar date. If these are defined to be the *x*-axis and *z*-axis respectively, then the pencils can be anchored on the base plane according to age at first event and date of first event, and the pencils will slope at 45 degrees to both the *x* and *z*-axes. If the *x*-axis is defined to be 'age at first event' and the *z*-axis to be 'date of first event' then *x* and *z* are constant over time, and the pencils will instead be vertical.

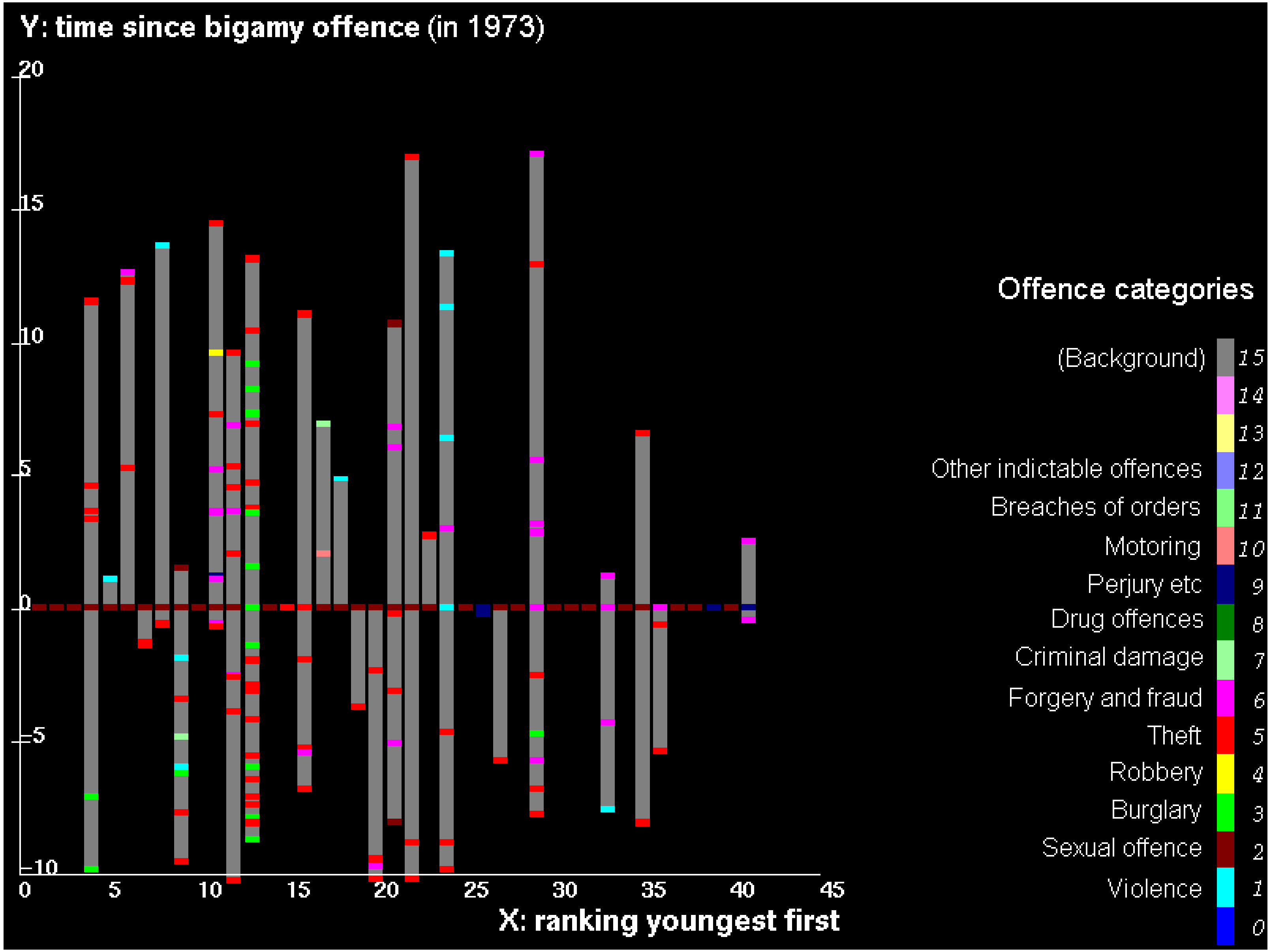


*Figure 3: A 2-D Lexis pencil diagram showing the criminal histories of 42 bigamy offenders who were convicted in England and Wales in 1973. The offenders are ranked with the youngest on the left and the oldest on the right. The representation of the criminal career starts with the first conviction and ends with the latest conviction in the study period. Each criminal conviction is represented by a bar of colour, with different colours representing the principal offence at that conviction.*

If there is no obvious second temporal dimension, then any other continuous variable can be used to define the *z*-axis and to position the pencils. One special useful case is using a temporal variable such as age for the *x*-axis and the rank order of age for the *z*-axis. This will neatly space out the pencils along a curve which represents the cumulative distribution function of age when viewed perpendicular to the *x*-*z* base plane.

The 3-D Lexis pencil display is a complex diagram, and a single view of the 3-D world will not be sufficient to explore a dataset. We developed these ideas using AVS (Advanced Visual Systems, 1992) , a scientific visualisation system which provides powerful user interactivity for exploring a three-dimensional world, such as zooming, panning and fly-through, as well as the programming environment to construct the Lexis pencils. Using the facilities provided in the developed AVS module, the user has control over the assignment of variables to axes, the number of faces on each pencil and the variables assigned to them, the colour mapping of codes to colours, the width of the pencil face and the angle subtended by adjacent faces. Users can also rotate and spin the pencils around their own axes while keeping the viewpoint fixed, and can display identification information such as case number for any pencil by clicking on it.

Another useful facility is case selection. Subsets of cases may be defined either by selecting cases of the values of certain variables (which may or may not be variables used to construct the display) or by specifying a set of case numbers. This allows users to compare, for example, males and females in separate displays without the necessity of increasing pencil complexity by adding an additional pencil face. Associated with this is the idea of transparency, where a subset of pencils can be made transparent and thus less intrusive on the eye, drawing attention to the fully opaque pencils.

### 5.1 Case study 1 continued - The Kirkcaldy families

We can now provide further information about the dataset of 188 Kirkcaldy married couples. One of the questions of interest in this study was the relationship between female employment and male unemployment. Discussion in the literature suggested that the wives of males who were unemployed were less likely to work because of the workings of the benefit system at that time - it was not financially advantageous for the female to work. This hypothesis was investigated by Davies et al (1991) who fitted a random effects logistic model to the probability of the wife working in any month. After controlling for the age of the youngest child, the number of children in the household, calendar time and the ages of both the husband and wife, an effect of the duration of the husband's unemployment was indeed found, with unemployment longer than one year being a significant deflator on the probability of the wife working.

Using the pencil construction outlined earlier, we can examine the data graphically using a 3D lexis pencil display, to see if we can find any further insights into the data. The pencils are arranged on an *x*-*z* base plane which is defined by age of the wife at marriage and date of marriage. We present two displays. With 188 pencils, we examine subsets of the data, and use the zooming facility of AVS to examine a small number of histories in detail. The first display (Figure 4) looks at couples who married in 1967 and 1968. Recall that the central face of the pencil is the employment history of the woman, with dark blue indicating unemployment, and the bottom face of the pencil is the age of the youngest child, with green indicating no children in the family, and yellow indicating a child aged less than one. What is clear is that women are stopping work before the birth of their first child - a pregnancy effect on the probability of a wife working. This effect is not surprising, but was omitted from the statistical analysis. Other subsets of marriage years show the same effect - the choice of marriage period is not significant.

The second display (Figure 5) shows those couples who married in 1952 and 1953. This display has made of the pencils semi-transparent, in order to highlight two pencils at the front of the display. The leftmost highlighted pencil is untypical for this time period, and illustrates a woman returning to work soon after childbirth - this is more typical of patterns in the 1970s. The rightmost highlighted pencil is more typical, and represents a woman who worked for a short period after marrying but stopped work a long time before the birth of her first child, and then stayed unemployed. The semi-transparent pencils in the background follow the same pattern as this leftmost pencil. In all displyed pencils the husband is working solidly with no evidence of unemployment. The general pattern of women's employment in the 1950s is different from that in the late 1960s and 1970s, where women more typically return to work around five years after their family is complete. This analysis suggests that there may be strong marriage cohort effects in the data, and again, no such effect was controlled for in the original analysis.

## 6. Difficulties encountered

We started this project by using AVS as our visualisation engine. This has a modular structure, and it was initially thought that programming the idea of the Lexis pencil display using the interactive module builder would be a relatively simple task. However, it soon became apparent that AVS had been designed for specific real world applications such as computational fluid dynamics and medical image processing. It became clear to us that AVS can still be used for our application, but a complex C program needed to be written both to read the event history data and to process it, providing such facilities as case selection, individual rotation of pencils and assignment of variables to faces which the standard AVS modules could not provide. This was converted into an AVS module and is available for downloading from the AVS repository and our own web pages.

Another disappointment was the ease of use of the AVS interface for panning and zooming. We used AVS on a Hewlett Packard workstation, using an X-terminal emulator running on a Pentium PC as the display. While speed of response was occasionally poor, a greater surprise was the paucity of user control for manipulating the viewpoint around the 3-D pencil world. Even with experience of the software, it was hard to control the viewpoint. In addition, it proved to be impossible to 'bookmark' a viewpoint once a suitable view had been reached. Most VRML viewers provide far better control , with options such as 'fly to', 'walk' and 'examine' which provide users with finer navigational control.

Finally, we have considered carefully the problem of dissemination. The software allows Postscript™ bitmaps of any view of the Lexis pencil world to be produced, and these can be published. However, it is often important to allow other users the ability to view the dataset themselves, and to identify their own special features of the data. We therefore provide the facility to write VRML (Virtual Reality Markup Language) files which define the 3-D Lexis pencil display. A difficulty with this was that at that time, surprisingly, no suitable VRML translator was available in AVS, and we therefore needed to write our own translator.

VRML viewers can now be purchased cheaply or downloaded free of charge on the World Wide Web to run on any platform. Other users should be able to easily load a VRML Lexis pencil file into their VRML viewer and fly through the 3-D Lexis pencil world on their own desktop computer.


*Figure 4: A 3-D Lexis pencil display showing a subset of the Kirkcaldy work history data. Those marrying in 1967 and 1968 are selected. The pencils are aligned on a 3-D plane according to age of the wife at marriage and date of marriage. The three faces from top to bottom display the work history of the husband, the work history of the wife, and the age of the youngest child in the household. A stong pregnancy effect on the probability of the wife working can be observed.*

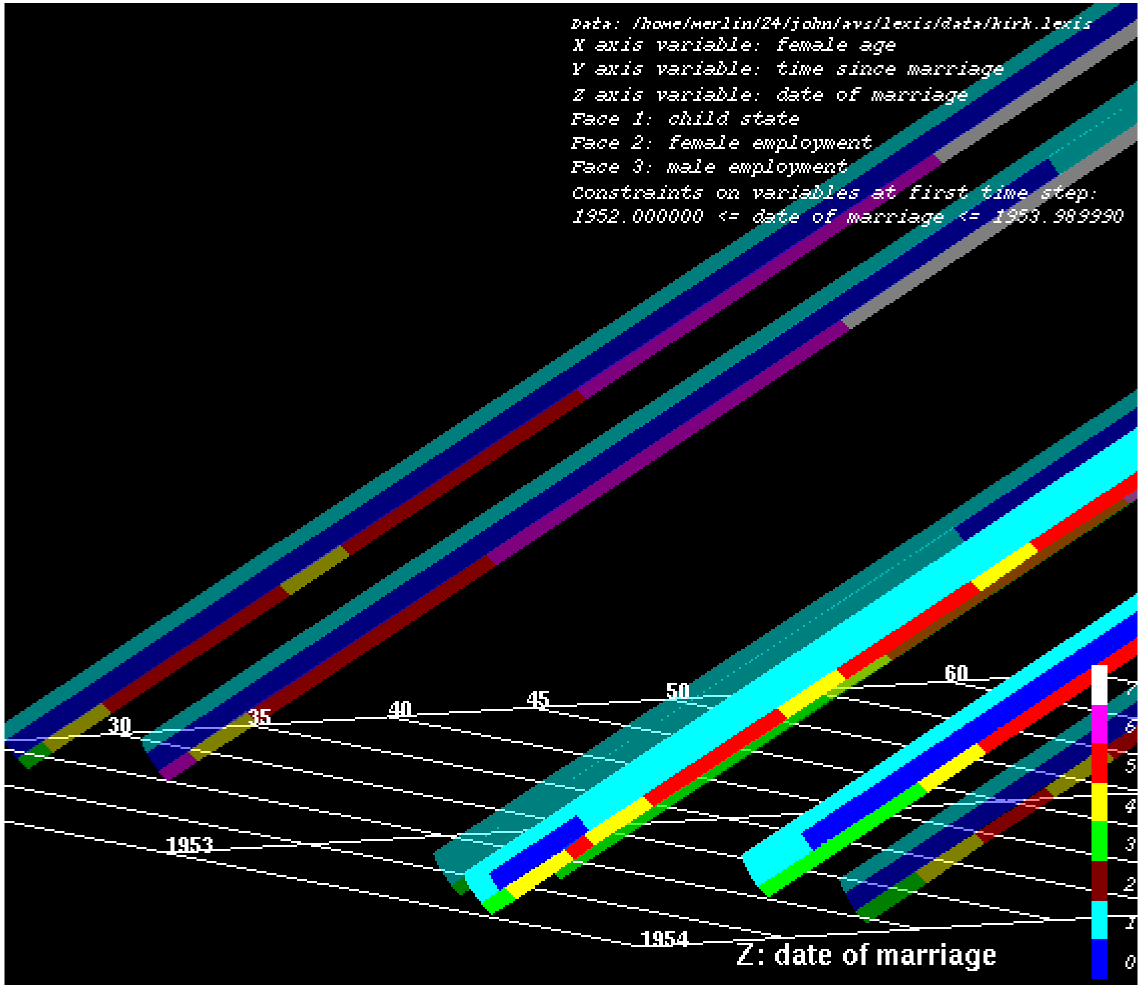


*Figure 5. Another view of the 3-D Lexis pencil display on the Kirkcaldy data. This display examines the subset of couples marrying in 1952 and 1953. The work history of the women follows a different pattern to those in later marriage cohorts. Some pencils are made semi-transparent to highlight the two histories in the foreground, which are discussed in the text.*

However, it appears that the complexity of the 3-D Lexis pencil world defeats many VRML viewers currently available. Under Windows 3.1, for example, Cosmoplayer and Live3D simply crash the browser, presumably because of lack of memory. VRML viewers appear to be better behaved on UNIX and Windows NT systems, although even a popular VRML browser such as VRweb fails to display labelling text in the Windows NT version. A robust, fully specified VRML viewer has yet to be found, but Cosmoplayer appears to us to give the best results.

Table 1 summaries what we see as the successes and failures of our approach.

| **Successes** | **Failures** |
|---|---|
| AVS software module written with full user control to read and process any event history dataset. | Visual complexity means that there are problems in examining more than 200 individual histories. |
| New substantive and statistical insights obtained into both datasets through the Lexis pencil display. | AVS complicated to use and substantial training needed. |
| Dissemination of ideas has been comprehensive, with published papers, international conference presentations and a web site. | Psychological issues of perception have not been addressed. |
| Pencil displays fit well into existing statistical literature. | |
| Dissemination of displays has been partially successful, with output provided as a static Postscript™ file or as VRML code. | |

**Table 1**
Evalutation of successes and failures of the case studies

## 7. The need for new tools

It is clear that most social scientists will not have access to AVS to use our AVS software module on their own datasets. AVS is an expensive software product and will not take priority in a list of essential software in the social sciences. Even if AVS is provided at their institution, it is not an easy package to use, and navigation around a 3-D world is not easy. We chose a scientific visualisation system originally as we wanted the ability to zoom, pan, and fly into a set of 3-D objects. Such facilities are now provided by VRML viewers and are far easier to use.

The next stage in this work is therefore to abandon AVS entirely as the visualisation engine.

Our AVS module currently reads event history data and processes it, before passing it to the AVS 'Geometry viewer'. This can be replicated by a window based user interface which will read and process the dataset, and which can generate VRML code. This code can be passed to a VRML viewer to provide the current user view. VRML viewers such as Cosmoplayer can already act as plugins for browsers such as Netscape, and could additionally plug in to this application. A language such as *tcl/tk* would be a good application generator and provides speedy development and customisable user interfaces in a variety of computing environments, including Windows, UNIX, and MacOS.

## Acknowledgments

The authors are grateful for the support of the UK Economic and Social Research Council (ESRC), who partially funded this research under the Analysis of Large and Complex Datasets initiative (grant number H 519255029). The 1973 sample of bigamists used as an example comes from another ESRC research project *Criminal Careers and Sex Offending* (grant number R 00023 6540), and we are grateful to Professor Keith Soothill for helpful discussions relating to this dataset. The sample of 188 married couples from Kirkcaldy comes from a third ESRC project funded under the Social Change and Economic Life initiative.

## References

Advanced Visual Systems (1992) *AVS Users Guide*, AVS Inc., Waltham, Mass.

Barry, J. T., Walby, S. and Francis, B. (1990) Graphical exploration of work history data. *Quad. Statist. Mat. Appl. Sci. Econ Socl*., **12**, 65-74

Blossfeld, H-P., Hamerle, A. and Mayer, K.U.(1989) *Event History Analysis*. Lawrence Erlbaum; New Jersey.

Cleveland, W.S. (1994) *The elements of graphing data*. Hobart Press, New Jersey.

Davies, R.B., Elias, P. and Penn, R. (1992) The relationship between a husband's unemployment and his wife's participation in the labour force. *Oxford Bulletin of Economics and Statistics*, **54**, 2, 145-171

Francis, B and Fuller, M (1996) Visualisation of event histories. *Journal of the Royal Statistical Society Series A*. **159**, 2, 301-8

Keiding, N. (1990) Statistical inference in the Lexis diagram. *Phil. Trans. R. Soc. Lond. A*, **332**, 487-509

Lexis, W. (1875) *Einleitung in der Theorie der Bevölkerungsstatistik*. Trübner: Strassburg

Pressat, R. (1961) *L'Analyse Démographique*. Presses Universitaires de France: Paris

Soothill, K., Sanderson, B., Peelo, M. and Ackerley, E. (1998) Bigamy - Likelihood of reconviction and the place of bigamy in the pantheon of crime. Submitted to *Medicine, Science and the Law*