WORKING PAPER

The LEXIS Computer Program for Creating Shaded Contour Maps of Demographic Surfaces

Bradley A. Gambill
James W. Vaupel
Anatoli I Yashin

August 1986
WP-86-037

International Institute
for Applied Systems Analysis
A-2361 Laxenburg/Austria
IIASA

The LEXIS Computer Program for Creating Shaded Contour Maps of Demographic Surfaces

Bradley A. Gambill
James W. Vaupel
Anatoli I. Yashin

August 1986
WP-86-37

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

The LEXIS Computer Program for Creating Shaded Contour Maps of Demographic Surfaces

Bradley A. Gambill, James W. Vaupel**, Anatoli I. Yashin****

INTRODUCTION

The LEXIS computer program, which was developed at the International Institute for Applied Systems Analysis (IIASA) and Duke University, is intended to aid demographers in the analysis of large arrays of data. Its application as a supplement to other methods of graphic display is demonstrated in *Thousands of Data at a Glance: Shaded Contour Maps of Population Surfaces* (Vaupel, Gambill, and Yashin, forthcoming) and will not be discussed here. This paper provides instructions on the use of the program, gives some hints concerning the art and craft of using the program in a creative way, and briefly describes the algorithm used in designing the program. A diskette containing a copy of the LEXIS program is enclosed. The program is copyrighted but the diskette is not protected against copying: please feel free to make and distribute copies. By making the program available to demographers and others interested in mapping the contours of surfaces, we hope to encourage the development of this method of data analysis. We would, of course, sincerely appreciate it if we and the International Institute for Applied Systems Analysis were acknowledged when the program or some modified version of it is used to produce maps for presentation or publication. Comments and suggestions are welcome!

*Bradley A. Gambill, T. Rowe Price Association, 100 East Pratt Street, Baltimore, Maryland 21202, USA
**James W. Vaupel, Humphrey Institute of Public Affairs, University of Minnesota, 301 19th Avenue South, Minneapolis, MN 55455, USA
***Anatoli I. Yashin, Population Program, IIASA, A-2361 Laxenburg, Austria

ADDRESS COMMENTS TO: Bradley A. Gambill, T. Rowe Price Association, 100 East Pratt Street, Baltimore, Maryland 21202, USA or to James W. Vaupel, Humphrey Institute of Public Affairs, University of Minnesota, 301 19th Avenue South, Minneapolis, Minnesota 55455, USA.

ADDRESS REQUESTS FOR ADDITIONAL OR UPDATED COPIES OF THE PROGRAM TO: either of the above or to Information Services, Population Program, International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria.

USE OF THE PROGRAM

Important Preliminaries

It will be assumed throughout this paper that the reader has some knowledge of the IBM personal computer and the DOS operating system. If this is not the case, we recommend that a potential user of the LEXIS program review the DOS manual provided with his or her IBM or IBM-compatible personal computer. Because the program does very little input/output checking, such knowledge may be quite helpful.

The diskette accompanying the paper contains two files: LX.COM and USFERT. The use of the program will be discussed shortly, but it is important that the program be copied to a diskette formated by the user's computer to include the DOS system for convenience. We recommend that the original diskette and a backup copy be stored in a safe place.

Printing Maps from the Screen

To print a LEXIS map from the computer screen, you must enter GRAPHICS (using your original DOS diskette) at the > DOS prompt before beginning the LEXIS program. Once the surface has been completely drawn on the screen, enter the <shift> and <PrtSc> keys simultaneously to print the map.

Format of Input Data

The data to be used by the LEXIS program must be in ASCII format, with m repetitions of a pattern consisting of the value of a label (such as the current year) followed by n data items (such as mortality rates at various ages). The value of n must be less than 150 and the value of m must be less than 300. For example, to plot a surface of mortality rates $q(a,t)$ from age 0 to 99 and year 1900 to 1982 the first data entry in the input file would be 1900, the second entry q(0,1900), the third entry q(1,1900), and so on up to the 101st entry which would be q(99,1900). The 102nd entry would be 1901, followed by q(0,1901), and so on up through q(99,1901). This pattern would be repeated up through the year 1982; the very last entry being q(99,1982). The label does not have to be a year: it could be an age or a life-expectancy or the value of any other variable that describes the data. Similarly, the successive data values following the label do not have to pertain to successive years of age, but could pertain to successive values of any other variable that describes the second dimension of the data. In any case, with input data in this format, the map will have the n-variable (e.g., age) on the y-axis, the m-variable (e.g., year) on the x-axis, and the data values (e.g., the q's) as the third dimension represented in the map.

Zeroes may be used to represent missing values in the data, if the actual data range does not include zero. This use of zeros is discussed below.

Beginning the LEXIS Program

To begin the LEXIS program, simply type LX at the DOS prompt. A title screen will appear at the start of the program and disappear after a few seconds, leaving the Main Menu on the screen.

Main Menu Options

A: Initializing data files

All data to be mapped must first be initialized by the LEXIS program. The procedure consists simply of entering the dimensions of the data file, and, optionally, searching the data for its highest and lowest points. This procedure begins with the selection of option *A* from the Main Menu. The user is prompted to enter the name and disk drive of the input data file, as well as the disk drive of the output file. The initialization process creates a file with the same prefix as the input file, and extension *.SPF*. The information contained in the .SPF file is provided by the user of the LEXIS program, as described in the next step below. A list of the required information will appear on the screen, and the cursor will move from place to place to prompt the user to enter the appropriate numbers. The screen will appear as follows (with illustrative numbers in bold print).

Notice that the x-axis information must be entered first, then the y-axis information. In our example of a mortality surface, we would enter the year-axis information, followed by the age-axis information. Once all necessary information has been entered, the program will ask if the data should be searched for the high and low. If y is entered, LEXIS will print the label (e.g., year) of the first column on the screen, followed by the label of each successive column. Checking the column labels as they appear on the screen is an excellent way to be sure that the data is in the proper format, and contains the proper number of entries in each data column. Once the initialized file is created, it is stored on the diskette and is available for all future mapping.

B: Selecting Display Colors

Selection of display colors is optional, with background color zero, line color zero, and a standard screen pattern being the default settings. An experienced user of the LEXIS program will find the selection of display colors to be a powerful option, however. By selecting this option, the user is able to select background color, line color, and pattern to be used in creating the map. The choice of background color is self-explanatory, and is a matter of personal taste for the user. The selection of line color depends on the data and the user's goals: if the lines are an important part of the map, they should be drawn in a color different than those used to create the patterns displayed (yellow for the print patterns, background for the screen patterns). If the exact location of the lines is less important than the general trends of change within the surface, then it may be more informative to hide the lines by making them the color of the most dominant color of the patterns being used. Experimentation will allow the user to determine which combinations of color will provide the most effective presentation of data.

The selection of pattern is a very powerful option provided by the LEXIS program. Included in the program are two sets of color combination patterns or "palettes" — the "screen" patterns which fade from green to yellow to red, and the "print" patterns which use the colors necessary to produce appropriate shades of gray when a contour map is printed on an IBM or Epson printer. The screen patterns are probably most pleasing to the eye, and provide distinguishable shades of logically progressing colors on a color screen; the print patterns provide the most easily distinguishable patterns when displayed on the printer or most monochrome monitors. It is possible to alter these patterns by entering 3 for pattern color and then placing a 1, 2 or 3 next to the appropriate region when prompted to do so. We have found it useful to shade ratio surfaces with just three colors, one representing the area below 1, one the region approximately equal to 1, and the last the region above 1. Other examples of use of this option are included in Vaupel, Gambill and Yashin, (1985); experimentation will help a user decide which color patterns lead to the most informative maps.

C: Creating a map from a data file

Once all preliminary initializations and selections are completed, a map can be created by selecting the *C* option from the Main Menu. The surface creation menu will appear and prompt the user chose a single data set map, a map developed from taking the ratio of two data sets, or a map created by taking the difference of two data sets. Once a selection is made, a new menu appears on the screen :

<F1> Files

To enter the name(s) of the input file(s) select the <F1> option when the bottom screen menu is highlighted. Since only one file is to be used in creating the map, only one input prompt will be given. If two files are to be used, the initial prompt will be for the numerator or minuend filename, and the second prompt for the denominator or subtrahend file for ratio and difference surfaces, respectively.

If the map will include only one data set, the user will be asked if the actual data include zeroes. If the data range includes zero, then "Y" should be entered at this prompt. If, however, zeroes are used only to denote missing values, in cohort data for example, then the user should enter "N" or <return>. For ratio surfaces, division by zero will be set equal to the background color.

<F2> Specs

The next step in creating a map is to enter the size of the surface to be displayed. The program will display the file specifications of each input file used in creating the surface, and then prompt the user to enter specifications for the map to be created. If the entire data set is to be mapped, it is not necessary to select this option. For comparative purposes, the dimensions may be extended beyond the borders of the data set in the x direction. In this case the missing data values will be assigned the background color.

<F3> *Smoothing*

Next, the user will be asked if the surface should be smoothed. If smoothing is desired, the user must enter an odd number less than or equal to twenty-one. The smoothing procedure simply replaces each point in the surface with the weighted average of the points in the n by n square surrounding the point. For example, if n is five, 25 points surrounding the original point will be averaged using the following weights :

1	4	6	4	1
4	16	24	16	4
6	24	36	24	6
4	16	24	16	4
1	4	6	4	1

When a point is near a boundary, the weighted average is taken over all those points within the surrounding square for which data are available.

<F4> *Lines*

Line selection is perhaps the most important decision that must be made in creating a shaded contour map. The LEXIS program provides seven different options for selecting the best lines:

<1> *Mortality* - Mortality surface lines begin with .000667 and then increase by 50% from one line to the next. We found these intervals to be useful in drawing contour maps of mortality surfaces. The contours are placed at the following levels:

B1 = 0.000667	B5 = 0.003377	B9 = 0.017094	B13 = 0.086541
B2 = 0.001001	B6 = 0.005065	B10 = 0.025642	B14 = 0.129811
B3 = 0.001501	B7 = 0.007598	B11 = 0.038463	B15 = 0.194717
B4 = 0.002251	B8 = 0.011396	B12 = 0.057694	

<2> *Fertility* - Fertility surface lines are placed at intervals we found to be convenient when analyzing fertility surfaces:

B1 = 0.001	B5 = 0.05	B9 = 0.13	B13 = 0.21
B2 = 0.01	B6 = 0.07	B10 = 0.15	B14 = 0.23
B3 = 0.02	B7 = 0.09	B11 = 0.17	B15 = 0.25
B4 = 0.03	B8 = 0.11	B12 = 0.19	

<3> *Ratio* - Ratio lines are placed at levels we sometimes found useful when analyzing surfaces that represented the ratio of two surfaces (e.g., male vs. female mortality). Each contour is 10 percent greater than the previous contour, as follows:

B1 = 0.513158	B5 = 0.751315	B9 = 1.1	B13 = 1.61051
B2 = 0.564474	B6 = 0.826446	B10 = 1.21	B14 = 1.771516
B3 = 0.620921	B7 = 0.0909091	B11 = 1.331	B15 = 1.948717
B4 = 0.683013	B8 = 1	B12 = 1.4641	

<4> *Difference* - The difference lines option allows the user to enter high and low values for a surface. The program then calculates the levels of 15 lines **between** the two points. These lines, depending on the user's choice, are placed either at even multiples or even intervals.

<5> *Multiples* - Even multiple lines option allows the user to enter a beginning point, and a multiplier. The first line is placed at the beginning point and the other lines are spaced at even multiples of the multiplier given. Thus, if 1.1 is the given multiplier, the lines increase by 10 percent, starting at the beginning point.

<6> *Auto-Select* - The auto select lines option takes the high and the low of the input file, and places 15 lines between the two points. The lines are placed at even multiples or even intervals, depending on the user's response to a prompt.

<7> *Selected* - The selected lines option allows the user to place the lines at any desired locations.

Once the lines have been chosen, enter *<8> Exit* to return to the menu at the bottom of the screen.

<F5> Menu

Entering function key 5 will return the user to the main menu of the LEXIS program.

<F6> Map

After all necessary information has been entered, the <F6> key can be used to start producing the map on the screen. The process can be interrupted only by re-booting the system. The last step in the mapping routine is the drawing of a border around the map. Once the map is completed, hitting any function key will put the user in the surface manipulation mode.

Surface Manipulation

When all calculations for the surface have been completed (drawing the border is the last step in the process), the user may view labeling and other pertinent information by, entering <F1>. A new screen will appear with the pertinent information, and a new screen manipulation menu at the bottom.

<F1> View Map - This option allows the user to flip back and forth between the map and information screens with the touch of a key.

<F2> Add Grid - This option enables the user to add a grid to the screen map in order to get more precise x and y axis readings. Within this option the user may chose where to start the grid on each axis, the grid line increments, the color of the grid lines, and also a color to "exclude" from the grid. We have found that it is sometimes more informative to leave the contour lines intact, on top of the grid, because it relegates the grid to the background, rather than to the foreground of the map. By entering <return> for each option, the default numbers listed on the screen will automatically be chosen.

<F3> Save Screen Map - This option allows the user to save the screen image on the output diskette for later viewing with the Recall procedure included in the Main Menu.

<F4> View Key - This option will put the color key of specified size on the screen for viewing or printing.

<F5> Return to Main Menu - This option ends the map manipulation procedure, and puts the user back at the main menu level.

D: Recall screen file

As mentioned in the previous section, it is possible to save a map with the <F3> command, and then recall it from a diskette with this command. After the file name and disk drive is entered, the file will appear on the screen, and all map manipulation options mentioned in the previous section will again be possible.

E: Return to DOS

Selecting this option will end the LEXIS session and return the user to the DOS operating system.

ALGORITHM AND DESIGN DECISIONS

To more fully understand the shaded contour maps created by the LEXIS program, the user may find it helpful to understand the general nature of the algorithm used in LEXIS and to be aware of certain design decisions that we made over the course of the program's development. The program resulted from frequent interaction between a demographer with some knowledge of computers and a computer science student with a keen interest in demography. Working together to develop the program and simultaneously conducting some substantive research with it enabled us to make substantial improvements in the program in its eighteen months of development. This section outlines some of the more subtle features of the shaded contour maps created by the program.

It is important to first discuss the differences that exist between the LEXIS contour maps and the more conventional contour maps often used by geographers, architects, geologists, and others interested in studying surfaces. The most obvious difference is that we have chosen to shade the LEXIS maps to show the various levels of the surface, rather than to label the lines with numbers. This procedure allows us to indicate "cliffs" on a surface without drawing multiple contour lines closely spaced together. That is, sudden surface changes of more than one level are represented by a single contour line, with the rapid change in level indicated by a jump in color or shade between the adjacent regions.

Because the LEXIS program does no interpolation before creating a surface, the appearance of a contour map depends on the number of data points used in creating it. Maps created with few data points will tend to include strong rectangular patterns, whereas contours on maps made from larger arrays of data will appear more rounded. It is possible to change the appearance of the maps with a little creativity. Rounded maps will become more rectangular if enlarged and viewed in portions. Rectangular contours can be smoothed by interpolating the data to produce a larger array of data points.

The LEXIS program determines the location of contour lines by replacing actual data points with integers representing the tier, or region of height, into which the data point falls. With fifteen contour lines, there are sixteen tiers on the surface. Next, each point is evaluated in terms of its relationship to adjacent points. Various patterns among nine-point squares are included in the program, and each point, with the eight points that surround it, is checked for conformity with one of these patterns. If the actual pattern is so complex that there is more than one discontinuity, the middle point is surrounded by a box. We chose this simple method of drawing lines to be a reasonable compromise between the competing goals of minimizing computational time and maximizing the smoothness and accuracy of contour lines.

Each data point in an array is assumed by the LEXIS program to represent the height of the surface in a box surrounding the point. For data pertaining to single years of time and age, each box is one year by one year. The tick marks placed along the edges of the map correspond to the mid-points of the boxes that they describe. Thus, a tick mark at age zero, say, will appear offset slightly from the origin and a tick mark at year 1980, in a map that runs through 1980, will appear slightly before the end of the horizontal axis.

If you have any questions or suggestions concerning LEXIS, please write to one of us at the addresses given at the beginning of this paper.

APPENDIX

Example Data Map

In order to alleviate the frustration that often results from using an unfamiliar computer program, this appendix will outline the procedure for creating a LEXIS map of the U.S. fertility data contained on the LEXIS distribution diskette. We will, however, only include necessary steps, and leave some of the most useful options for users to discover for themselves.

Step 1: Beginning the LEXIS program

Before beginning the procedure, create a work diskette by copying all of the files contained on the distribution diskette to a diskette formatted by your machine. Put this diskette into drive a: (the left side diskette drive).

Step 2: File Initialization

The data contained on the diskette is U.S. fertility rates (Heuser, 1984) and runs from age of mother 14 to 49 and year 1945 to 1980. To initialize this data, enter *A* from the main menu. Information should be entered as follows:

```
                     File Initialization
                     -------------------
 Input File Name: USFERT________

    Disk Drive: A:                 Output File Disk Drive: A:

                     FILE SPECIFICATIONS
                     ===================
                  X-Axis Information
                  --------------------
                  Starting x:   1945
                    Ending x:   1980
                      Label :   YEAR

                  Y-Axis Information
                  --------------------
                  Starting y:   14
                    Ending y:   49
                      Label :   AGE

          Search for surface High and Low ?  Y
```

Once this has been completed, enter *y* to search the data set. The year numbers will appear on the screen as the program reads the data, beginning with 1945 and ending with 1980. After the file has been completely initialized, the main menu will reappear.

Step 3: Creating the Map

The color pattern is set to the screen patterns by default, so we will ignore Main Menu option *B* and select *C* to create the LEXIS map. The Map Creation menu will now appear on the screen, giving the user the option to create a LEXIS map using one or two input files. Since there is only one data file on the diskette, enter *A*.

The appropriate information should be entered so that the screen looks as follows before continuing:

```
  Input Files:             Drive   | Specifications:
  USFERT______              A      |      Start YEAR:          1945
                                   |      End YEAR:            1980
  ____________              _      |      Start AGE            14__
Zero in Data Range (Y/N)?   Y      |      End AGE              49__
-----------------------------------------------------------------------
 Line Options:           |                   | L1:   0.001000  (low)
                         |                   | L2:   0.010000
   <1> Mortality         |                   | L3:   0.020000
   <2> Fertility         |                   | L4:   0.030000
   <3> Ratio             |                   | L5:   0.050000
   <4> Difference        |                   | L6:   0.070000
   <5> Multiples         |                   | L7:   0.090000
   <6> Auto-Select       |                   | L8:   0.110000
   <7> Selected          |                   | L9:   0.130000
   <8> Exit              |                   | L10:  0.150000
                         |-------------------| L11:  0.170000
                         |   High:           | L12:  0.190000
                         |     0.258900      | L13:  0.210000
                         |   Low:            | L14:  0.230000
                         |     0.000010      | L15:  0.250000  (high)
-----------------------------------------------------------------------
 <F1> Files   <F2> Specs   <F3> Smoothing   <F4> Lines   <F5> Menu   <F6> Map
```

To input the appropriate information, first enter *<F1>*. You will prompted to enter the filename to which you must respond *USFERT <return>*, and then the disk drive of the input file. Enter the appropriate letter here and <return>. You will then be asked if zero is in the data range. Since U.S. fertility rates could be 0 at one the ages and times covered, enter *y*. Next, input <*F*4> to select the appropriate contour lines. Since we are mapping U.S. fertility, enter 2, and then 8 to return to the screen menu. Since we want to view the entire surface, there is no need to use the <F2> option.

Now, enter <*F*6> to start creating the map. The process will take about one minute to complete, and will be entirely finished when a yellow border appears around the map.

Step 4: Screen Manipulation and ending

After the entire map has been generated, entering the <*F*1> key will reveal a text screen with labeling information and the screen manipulation menu. Also, on the last line will appear a screen manipulation menu which outlines user options at this point in the mapping process. In order to see the color key used in creating the map, enter <*F*4>. After the color key appears on the screen, hitting <return> will return the user to the information screen. From here, enter <*F*5> to end the process and return to the Main Menu.

Following this example should give the user a practical, but superficial, understanding of the LEXIS program capabilities. It is important that this example be used only as a supplement to, rather than substitute for, the more extensive description contained in the first part of the paper.

REFERENCES

Heuser, R.L. (1976) Fertility tables for birth cohorts by color: United States, 1917-1973. U.S. Department of H.E.W., Health Resources Administration, Rockville, MD. DHEW Publication No. (HRS) 76-1152.

Vaupel, J.W., B.A. Gambill, and A.I. Yashin (1986) ***Thousands of Data at a Glance: Shaded Contour Maps of Demographic Surfaces***. Forthcoming Research Report. Laxenburg, Austria: International Institute for Applied Systems Analysis.