APPENDIX

This Appendix provides some programs to do bootstrapping. They depend on EXCEL macros, but users do not need to know details of the programming. All one needs to do is insert the data in a spreadsheet and run the macro. This is accomplished by opening the “TOOLS” menu, and selecting “MACRO” from the submenu. A window labeled “Macro” comes up, and hitting the “run” button starts the program. A keystroke alternative is available on computers running Windows (in some machines it is ALT F8). This brings up the same “Macro” window, and is a little faster than going thru the TOOLS menu. When you open one of these programs, a window may come up indicating that the program contains macros, and warning about viruses. Hit “Enable Macros” to use the program. Some virus protection software will also detect the code used and bring up a warning window.

GENERAL PURPOSE PROGRAMS

There are 4 programs available to produce the results of Chapters 2 and 3. The first, BOOT1, bootstraps single columns of data. It computes the bootstrapping example of Section 2.1. Contents of the EXCEL window are as follows:

BOOT1.xls

	A	B	C	D	E
1	NOBS	NBOOT	BOOTSTRAP	69.1	PARAM
2	10	30	SAMPLE	139.4	INPUT
3	DATA			101	66
4	13	12	11	121.2	
5	106	13	301	86.6	
6	203	9	8	142.4	
7	131	5	106	88.1	
8	160	9	8	96.7	
9	8	4	13	107.9	
10	67	12	11	103.8	
11	61	10	67	112.2	
12	11	7	131	105.3	
13	301	9	8	64.2	
14		AVE	66.4	96.1	
15				90	
16				112.5	
17				136.9	
18				122.3	
19				125.6	
20				137	
21				82.5	
22				91.3	
23				109.7	
24				83	
25				110.7	
26				118.2	
27				156.6	
28				99.9	
29				95.1	
30				66.4	

Sheet1 Sheet2

The 5 left-hand columns of the worksheet should be reserved for the program. Various calculations can be carried out in the remainder of the spreadsheet, and will ordinarily be used to perform intermediate calculations in bootstrapping, as described below. To modify the program, one inserts the data in the lefthand column, the number of observations under NOBS and the number of bootstraps to be run under NBOOT. The actual bootstrap results appear in the 4^{th} column which usually will have 1,000 to 2,000 entries (so nothing should be entered under this column in the worksheet, and one needs to remember to print out only the first page as simply hitting “print” may get you 40 pages or so!). The program presently is set up to provide bootstraps of the average as in Section 2.1.

Ordinarily there is little reason to bootstrap averages, and this worksheet is used only to show how to use the program and to correspond to the example of Section 2.1. The first column contains the actual observations so NOBS has to conform to the number of observations entered (10 here). Under NBOOT, enter the number of bootstraps you want to run. Ordinarily this will be 1,000 or so, but only 30 are used for illustration. The second column will show the serial number of the observation randomly selected by the program, and the third column shows the bootstrap sample (or “bootstrap replication”). Run the program and you will see these numbers change as each sample is drawn. Any function you want to bootstrap needs to be inserted somewhere. Presently it is in the third column below the bootstrap sample. In the present example, the average is calculated there. For convenience in keeping track of things, a label “AVE” has been inserted to the left of the average calculated from the bootstrap sample. Note, however, that if you use a larger number of observations, the program will write over the word “AVE”, and over the calculated value of the average, so you have to be careful to remove both “AVE” and the calculation to the right of that word before you change the number of observations.
In general, be careful what you do in the left-hand 5 columns or you may get some strange results. It is wise to save copies of the program as you experiment with changes, so that you can go back if something seems to be wrong with the current version.

The actual input to the program is in cell E3 (shaded in the worksheet, and just below “PARAM INPUT”) and if you highlight that cell you will see that it is directed to cell C14, which contains the function =AVERAGE (C4:C13) which averages the bootstrap replication. You can remove this cell and substitute another function. Try it with the variance function, =VAR(C4:C13) and run the program. It should produce 30 bootstrap values of the variance in column 4. The basic idea here is that you can compute anything that can be done directly on the EXCEL spreadsheet and enter it in cell E4 (“PARAM INPUT”) and 30 bootstrap values will be produced in column 4. Change NBOOT to 1000, and run the program. Then order the entries in column using the SORT function in the DATA window, and locate entries 25 and 975. These are the percentile confidence limits as described in Section 2.5.

Suppose you want the bootstrap standard error of eq.(2.1) for variances computed as above. The equation is:

$$\hat{se}_{boot} = \left\{ \frac{\Sigma[s(x^{*b}) - s(\cdot)]^2}{B-1} \right\}^{1/2} \qquad (2.1)$$

Find the bootstrap mean (average of the bootstrap values in column D; this is s(·) in eq. (2.1)), and the calculation is shown in the following table:

NOBS	NBOOT	BOOTSTRAP	12861.2	PARAM
10	30	SAMPLE	14468	INPUT
DATA			12173.7	6754.84444
13	13	301	7102.1	
106	11	61	9249.78	
203	8	160	4737.57	
131	10	67	8640.27	
160	11	61	10479.2	
8	5	106	8991.57	
67	9	8	6107.61	
61	10	67	5704.9	
11	7	131	14642.4	
301	8	160	6204.9	
	VAR	6754.8444	7834.23	
			13931.7	
			2359.88	
			4403.33	
			16701.6	
			8079.73	
			8469.73	
			4544.71	
			6070.67	
			4557.6	
			5182.44	
			5182	
			8124.77	
			13266.1	
			9682.27	
			5299.43	
			6754.84	
		MEAN	8393.61	

You can then conduct the calculation in column F by subtracting the overall mean from the individual values of variances [$s(x^{*b})$ in the equation above] and squaring (remember to use the $ notation to "freeze" the value of the bootstrap mean), sum up the column of squared values, divide by B-1 and find the square root. This is illustrated for the variance below [all we do is add a column to the table above , sum, divide by B-1 (30 - 1= 29 here) and take the square root]. Note that the additional calculations are simply added to the worksheet after the actual bootstrapping is done.

Bootstrap standard error calculation:

NOBS	NBOOT	BOOTSTRAP	12861.2	PARAM	19959076	
10	30	SAMPLE	14468.0	INPUT	36898367	
DATA			12173.7	6754.84444	14289344	
13	13	301	7102.1		1667994	
106	11	61	9249.8		733026	
203	8	160	4737.6		13366642	
131	10	67	8640.3		60840	
160	11	61	10479.2		4349738	
8	5	106	8991.6		357554	
67	9	8	6107.6		5225784	
61	10	67	5704.9		7229153	
11	7	131	14642.4		39047395	
301	8	160	6204.9		4790445	
	VAR	6754.84444	7834.2		312901	
			13931.7		30670826	
			2359.9		36405907	
			4403.3		15922296	
			16701.6		69022169	
			8079.7		98518	
			8469.7		5795	
			4544.7		14814011	
			6070.7		5396059	
			4557.6		14714961	
			5182.4		10311575	
			5182.0		10314429	
			8124.8		72276	
			13266.1		23741173	
			9682.3		1660640	
			5299.4		9573920	
			6754.8		2685548	
		MEAN	8393.6		393698361.5	SUM
					3684.536	BOOTSTRAP. STANDARD ERROR

In the above case we produced the bootstrap values and then operated on these to find the bootstrap standard error. In other situations one may want to insert several stages of calculations before doing the bootstrapping. Consider the "parametric regression" calculations of Section 2.6. One needs first to find the residuals about regression. This can be done separately or on the same worksheet. For convenience the regression calculation is listed separately here, as follows.

EXCEL calculation for a regression line:

ORIGINAL DATA

X	Y
10	12.672
12	8.9391
14	13.934
15	16.377
17	13.252
21	19.121
23	17.821
28	18.879
30	21.047
35	25.213

Regression Statistics	
Multiple R	0.9172596
R Square	0.8413651
Adjusted R Square	0.8215358
Standard Error	1.9880968
Observations	10

ANOVA

	df	*SS*	*MS*	*F*
Regression	1	167.70690	167.70690	42.4302
Residual	8	31.62023	3.95253	
Total	9	199.32714		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	6.15279	1.74062	3.53483	0.0076
X Variable 1	0.51574	0.07918	6.51385	0.0001

RESIDUAL OUTPUT

Observation	*Predicted Y*	*Residuals*
1	11.31021	1.36179
2	12.34170	-3.40260
3	13.37318	0.56082
4	13.88893	2.48807
5	14.92041	-1.66841
6	16.98338	2.13762
7	18.01487	-0.19387
8	20.59358	-1.71458
9	21.62506	-0.57806
10	24.20378	1.00922

Copy the residuals into the first column of BOOT1, and copy the “Predicted” values in a column to the right of the “active” part of the sheet (first 5 columns). Construct new Y values by adding the bootstrapped residual to the predicted values. Calculate a slope (using the SLOPE function) from the new Y values and the X-values and enter this in the PARAM INPUT box (one could calculate the slope directly in this box, but it may be best to calculate it elsewhere on the spreadsheet and set PARAM INPUT equal to this value. If you want to also bootstrap the intercepts, obtain means of the new Y and the X values and compute an intercept using the slope value. For simplicity, the present version of BOOT1 only handles one parameter, but one can run the program twice, tabulating slopes and then intercepts. The modified program follows. If you want intercept calculations, then add these and run the program again. One would usually also order the slope values and find confidence limits, and make a histogram to plot the frequency distribution. With a little effort all of this can appear on the same spreadsheet so you can have all of your results together, and then pick off what's needed for a report. Run the program as set up above to try it out and better understand the descriptions given here.

Regression bootstrap:

NOBS	NBOOT	BOOTSTRAP	0.5906	PARAM			
10	20	SAMPLE	0.5250	INPUT			
DATA			0.5118	0.57185669	PREDICTED	NEW Y	x
1.3618	5	-3.40260	0.5497		11.31021	7.90761	10
-3.4026	7	2.48807	0.4476		12.34170	14.82977	12
0.5608	7	2.48807	0.5756		13.37318	15.86126	14
2.4881	13	1.00922	0.4399		13.88893	14.89815	15
-1.6684	12	-0.57806	0.5550		14.92041	14.34235	17
2.1376	6	0.56082	0.5187		16.98338	17.54420	21
-0.1939	4	1.36179	0.4442		18.01487	19.37665	23
-1.7146	10	-0.19387	0.4936		20.59358	20.39971	28
-0.5781	7	2.48807	0.5645		21.62506	24.11314	30
1.0092	13	1.00922	0.5426		24.20378	25.21300	35
			0.5342				
			0.5004		SLOPE	0.57186	
			0.4737				
			0.4817				
			0.4568				
			0.4901				
			0.5719				

BOOT2

To use nonparametric bootstrapping for regression data one needs a program that handles 2 columns of data. This is done in BOOT2 where the first 2 columns are input into a program. The version of BOOT2 furnished is set up to bootstrap the same data set as used above. However, parametric regression bootstrapping should be used for smaller data sets (less than, say, 10 pairs of observations). One can see why this is so by running the program and plotting a frequency distribution of outcomes for smaller samples. In the example shown here parametric and nonparametric regression gave about the same frequency distributions. For a more realistic example we use the larger data set of Example 3.4 and bootstrap the variance about regression. The first table below shows BOOT2 as it is currently set up, while the second table gives it with the example used here (n = 19).

BOOT2.xls

	A	B	C	D	E	F	G
1	N0BS	NBOOT				0.525633	PARAM
2	10	20				0.580319	INPUT
3	DATA X	Y	RAND NO	X	Y	0.522876	0.61692
4	10	12.672	8	17	13.252	0.427133	
5	12	8.939	8	17	13.252	0.562192	
6	14	13.934	13	35	25.213	0.522399	
7	15	16.377	9	21	19.121	0.404017	
8	17	13.252	8	17	13.252	0.467963	
9	21	19.121	13	35	25.213	0.495395	
10	23	17.821	11	28	18.879	0.593733	
11	28	18.879	5	12	8.939	0.489994	
12	30	21.047	4	10	12.672	0.521930	
13	35	25.213	5	12	8.939	0.507591	
14			SLOPE	0.6169		0.595299	
15						0.516997	
16						0.451051	
17						0.501070	
18						0.560762	
19						0.569541	
20						0.616916	

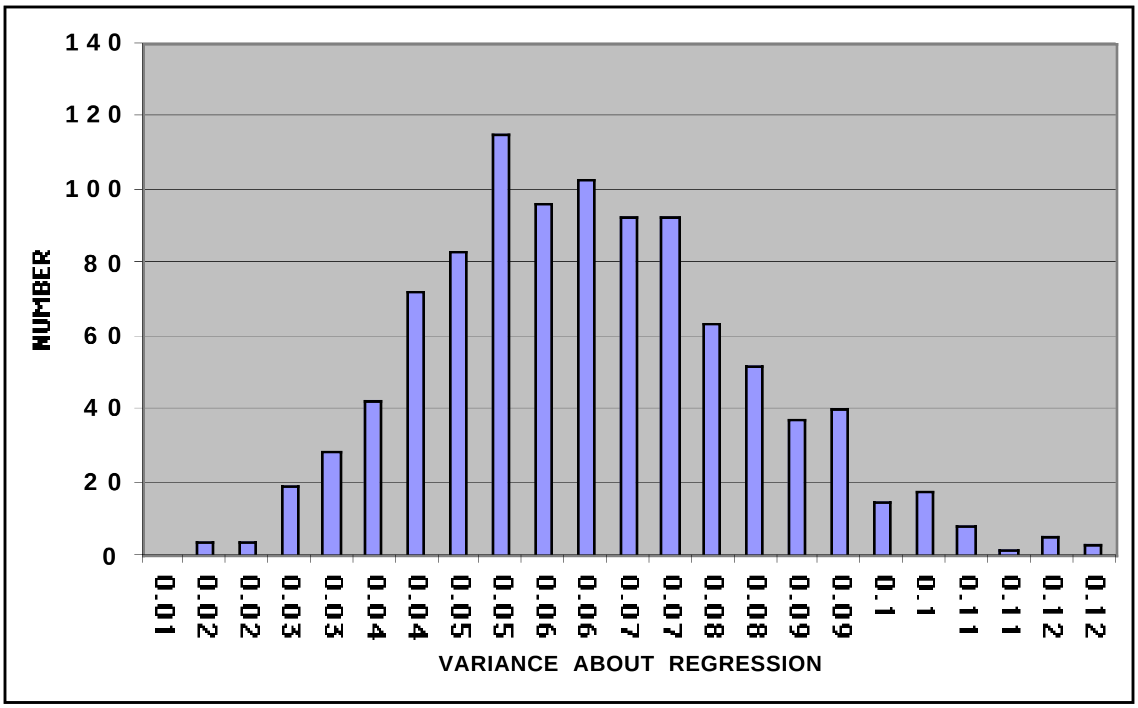
TWO COLUMN INPUT | Sheet2 | S

BOOT2 WITH BEAR DATA

N0BS	NBOOT				0.08994	PARAM		
19	2000				0.07248	INPUT	SUM OF	
DATA X	Y	RAND NO	X	Y	0.02706	0.04643	SQUARES	
1	2.8332	18	15	3.2189	0.09502		0.03354	
2	2.5649	16	13	2.9444	0.04895		0.00143	
3	2.1972	10	7	2.3979	0.04463		0.01468	
4	2.5649	15	12	2.5649	0.04909		0.07679	
5	2.4849	7	4	2.5649	0.04848		0.05740	
6	2.6391	20	17	3.1355	0.07566		0.00087	
7	2.3979	10	7	2.3979	0.04953		0.01468	
8	2.5649	19	16	3.1781	0.06649		0.00604	
9	2.8332	21	18	2.9957	0.10144		0.05468	
10	2.1972	18	15	3.2189	0.06239		0.03354	
11	3.2189	20	17	3.1355	0.05743		0.00087	
12	2.5649	10	7	2.3979	0.04019		0.01468	
13	2.9444	16	13	2.9444	0.06976		0.00143	
14	2.7726	16	13	2.9444	0.04449		0.00143	
15	3.2189	19	16	3.1781	0.05643		0.00604	
16	3.1781	13	10	2.1972	0.01956		0.26588	
17	3.1355	8	5	2.4849	0.05609		0.00903	
18	2.9957	14	11	3.2189	0.02939		0.19490	
19	2.9957	16	13	2.9444	0.05252		0.00143	
		MEANS	12.053	2.84541	0.06192		0.78934	SUM
		SLOPE	0.0646				0.04643	VARIANCE
		INTERCEPT	2.067					ESTIMATE

To bootstrap the variance about regression, one adds calculations for means, slope and intercept as above and uses these to calculate a sum of squares and variance estimate which is then loaded into PARAM INPUT. The bootstrap results (B=1000 here) can then be summarized by HISTOGRAM (not shown here) and a frequency diagram constructed with the CHART WIZARD. The frequency plot indicates that the estimate of the variance about regression is itself quite variable. The value for the original data set is 0.0645 while the average of 1,000 bootstraps is 0.0583, suggesting a bias [eq.(3.2)] of 0.0062 which is likely not of much concern, considering the variability of the estimate as displayed below.

FREQUENCY DISTRIBUTION OF VARIANCES ABOUT REGRESSION FOR BEAR DATA



The next two programs likely do not need much explanation. They serve to jackknife one (JACK1) and two-column (JACK2) data. Applying JACK1 to the data of Section 2.2 gives the result shown below. All one needs to do is load the data in the left hand column and enter the number of observations. Running the macro then produces the basic jackknife data set, and one can proceed to add whatever calculations are desired. JACK2 is used in the same way, and is demonstrated here with the bear data (only part of the 19 sets of output is shown).

JACK1.xls

	A	B	C	D	E	F	G	H	I	J	K	L
1	NOBS											
2	10											
3	DATA											
4	13		106	13	13	13	13	13	13	13	13	13
5	106		203	203	106	106	106	106	106	106	106	106
6	203		131	131	131	203	203	203	203	203	203	203
7	131		160	160	160	160	131	131	131	131	131	131
8	160		8	8	8	8	8	160	160	160	160	160
9	8		67	67	67	67	67	67	8	8	8	8
10	67		61	61	61	61	61	61	61	67	67	67
11	61		11	11	11	11	11	11	11	11	61	61
12	11		301	301	301	301	301	301	301	301	301	11
13	301											

Sheet1 / Sheet2 / Sheet3

JACK2.xls

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	NOBS												
2	19												
3	DATA												
4	1	2.8332		2	2.5649	1	2.8332	1	2.8332	1	2.8332	1	2.8332
5	2	2.5649		3	2.1972	3	2.1972	2	2.5649	2	2.5649	2	2.5649
6	3	2.1972		4	2.5649	4	2.5649	4	2.5649	3	2.1972	3	2.1972
7	4	2.5649		5	2.4849	5	2.4849	5	2.4849	5	2.4849	4	2.5649
8	5	2.4849		6	2.6391	6	2.6391	6	2.6391	6	2.6391	6	2.6391
9	6	2.6391		7	2.3979	7	2.3979	7	2.3979	7	2.3979	7	2.3979
10	7	2.3979		8	2.5649	8	2.5649	8	2.5649	8	2.5649	8	2.5649
11	8	2.5649		9	2.8332	9	2.8332	9	2.8332	9	2.8332	9	2.8332
12	9	2.8332		10	2.1972	10	2.1972	10	2.1972	10	2.1972	10	2.1972
13	10	2.1972		11	3.2189	11	3.2189	11	3.2189	11	3.2189	11	3.2189
14	11	3.2189		12	2.5649	12	2.5649	12	2.5649	12	2.5649	12	2.5649
15	12	2.5649		13	2.9444	13	2.9444	13	2.9444	13	2.9444	13	2.9444
16	13	2.9444		14	2.7726	14	2.7726	14	2.7726	14	2.7726	14	2.7726
17	14	2.7726		15	3.2189	15	3.2189	15	3.2189	15	3.2189	15	3.2189
18	15	3.2189		16	3.1781	16	3.1781	16	3.1781	16	3.1781	16	3.1781
19	16	3.1781		17	3.1355	17	3.1355	17	3.1355	17	3.1355	17	3.1355
20	17	3.1355		18	2.9957	18	2.9957	18	2.9957	18	2.9957	18	2.9957
21	18	2.9957		19	2.9957	19	2.9957	19	2.9957	19	2.9957	19	2.9957
22	19	2.9957											

JACK2 / Sheet2 / Sheet3

The two columns of input data are on the left and the first 6 jackknife results appear to the right, with the first pair left out, then the second, third and so on. One can add various calculations (e.g., slopes) directly on the spreadsheet. It is wise to clear out all but the input columns before running the programs, i.e., JACK1 should appear as shown below before you run it. Inasmuch as it only takes a few minutes to run most of the programs, it is advisable to make a second run starting from the original program if anything looks questionable in your results. When using 2000 bootstraps with some mildly complicated secondary calculations, such as BOOT2 with the variance about regression calculations as shown above, it does take some time for the program to run. On a MAC PowerBook G3 it took about 15 minutes, while a HP Pavilion portable (100mhz CPU) took about 10 min.

BASIC ARRANGEMENT FOR JACK 1

```
NOBS
   19
DATA
     1    2.8332
     2    2.5649
     3    2.1972
     4    2.5649
     5    2.4849
     6    2.6391
     7    2.3979
     8    2.5649
     9    2.8332
    10    2.1972
    11    3.2189
    12    2.5649
    13    2.9444
    14    2.7726
    15    3.2189
    16    3.1781
    17    3.1355
    18    2.9957
    19    2.9957
```

SPECIAL PURPOSE PROGRAMS USING BOOTSTRAPPING

The programs given above are general purpose in that they are written to use one or two columns of data, and the results can be modified on the spreadsheets to provide a variety of results. Some further programs follow, designed for more specialized purposes. However, the same underlying theme is pursued. That is, the MACROS bootstrap a data set, and the user makes changes on the EXCEL spreadsheet to accomplish a variety of different purposes.

Approximating the Lotka-Leslie model

The first special purpose program to be considered here is APPLMB, which has been prepared to bootstrap eq.(11.9), the approximation suggested by Eberhardt (1985) and further evaluated in Eberhardt (2002). The underlying equation is:

$$\lambda^{\mathbf{a}} - s\lambda^{\mathbf{a-1}} - l_{\mathbf{a}} m \left[1 - \left(\frac{s}{\lambda}\right)^{\mathbf{w-a+1}}\right] = 0 \qquad \mathbf{(11.9)}$$

An example of the program using data on manatees follows, with the equation associated with the box labeled RES shown.

PROGRAM APPLMB:

RES = =G5^E5-B5*G5^(E5-1)-D5*C5*B5^(E5-1)*(1-(B5/G5)^(F5-E5+1))

APPLMB.xls

	A	B	C	D	E	F	G	H	I	J	K
1											
2		NOBS	RATE								
3		9	1.01						Estimate		No. of
4		S	SJUV	m	a	w	lbd	RES	of lambda	Residuals	iterations
5		0.9632	0.8444	0.1665	4	50	1.0660	-0.00009			
6		0.9548	0.8889	0.1327					1.0437	-0.0001	105
7		0.9648	0.7556	0.1629					1.0571	-0.0001	112
8		0.9656	0.8667	0.1432					1.0584	-0.0001	107
9		0.9690	0.8222	0.1567					1.0649	-0.0001	118
10		0.9828	0.8667	0.1515					1.0804	-0.0001	125
11		0.9633	0.7778	0.1583					1.0555	-0.0001	113
12		0.9751	0.8000	0.1763					1.0784	-0.0001	123
13		0.9680	0.8000	0.1493					1.0580	-0.0001	117
14		0.9632	0.8444	0.1665					1.0660	-0.0001	109
15											
16											

The main program does not do any bootstrapping. All of the bootstrapping is done in advance by using BOOT1 and BOOT2, and is done on the three input variables, S, SJUV, and m before using APPLMB. The only purpose of APPLMB is to to produce estimates of lambda from the input data. One thus has to bootstrap adult survival data (S) and enter, say, 1,000 bootstrap values in that column. Values of m are similarly bootstrapped and entered in the column headed by m. The age of first reproduction (a) and the maximum age (w) are fixed and need only be entered once. SJUV corresponds to l_a in eq.(11.9). If survival is assumed to be at the adult rate after the first year of life then the first-year survival can be entered in the column headed SJUV (this was done for manatees as there were no estimates of early survival beyond age 1). In some examples, however, there are at least 2 early survival values, so SJUV has to be calculated as the product of two or more survival rates. Each of these rates has to be bootstrapped separately and the results combined (multiplied) to get SJUV, and the equation governing RES modified accordingly. Consequently, one will have several different bootstrapping sheets containing the data to be entered into APPLMB. In the example shown above there are only 9 entries just to show how the program is set up; in the usual case there will be at least 1,000 entries in the three left-hand columns.

The actual calculations are done by an equation in the spreadsheet associated with the box headed RES. The program runs thru all 1,000 rows, calculates lambda in the box headed lmb, and stores the estimates of lambda in the column headed "Estimate of lambda". The value of the residuals is also listed to the right, so that one can check to see that the program converged properly. A maximum of 2000 iterations is used to solve eq.(11.9), so if the last column contains the number 2000, the program likely did not converge adequately. Users need to understand that eq.(11.9) is nonlinear in nature, and the program may thus fail to converge. Hence a very important item appears in the box headed RATE. This is the initial "guesstimate" of lambda. I suggest setting NOBS at 30 (which limits the number of bootstraps used to the first 30), and running the program. If there are problems with convergence (i.e., 2000 shows up repeatedly in the box headed "Number of iterations"), then experiment with the value used for RATE until the number of iterations is less than 2000. EXCEL contains the very useful program, SOLVER, in the TOOLS menu. When you activate SOLVER, a box comes up asking for the “Target” cell, and this should be set as the box under RES. It also asks for the cell to vary and this is the box under “lbd”. When you run the program APPLMB, you will note that it sequentially loads all of the data into the first line and solves eq. (11.9) with this data. Hence if you want to use SOLVER to check a result you have to copy data corresponding to the case in question into the first 3 columns of the first row (columns headed S, SJUV, and m), and then activate SOLVER. Be sure to read the “help” items associated with SOLVER, as there are various things that you can do to facilitate getting a solution, and to check that the solution is appropriate. I have used the routine built into APPLMB on many data sets, and have not had any important problems, so long as I used good “guesstimates” to start with. You can demonstrate failure of the routine by putting in an unreasonable values of RATE – -it will yield all sorts of strange outcomes. Incidentally, λ should never be set equal to s. I usually start with $\lambda = 1.01$. If you get error messages, usually the box that comes up will have a button titled “END”, and pressing this gets you back to the spreadsheet.

For an example, we use the data on Yellowstone grizzly bears that produced the frequency distributions of Fig. 11.8. The data shown there were produced by a program written especially for the purpose. Using the programs given here reduces the effort required considerably and does not require a knowledge of a programming language. Applications of BOOT1 and BOOT2 produce the basic data. Survival can be estimated in various ways, but in this case radiotelemetry was used, so the individual records are those of “radio-days” of observation per individual bear. Some of the individuals died during the course of the study, and a simple estimator was used:

$$S = (1 - d/days)^{365}$$

Where d = number of deaths observed and days = number of “radio-days” tallied. The bootstrapping program contains a series of records of individual bears with deaths tallied as 1’s and survival tallied as 0’s in the y-column while the corresponding days of observation for the individual appear in the x-column. The data are as shown in the table below, and were inserted in BOOT2 or BOOT1 (for reproductive rate) and 1,000 bootstrap samples drawn. Bootstrap values (means of 1,000 bootstraps) are shown in the table along with the direct estimates of the several parameters.

INPUT DATA FOR PROGRAM APPLMB APPLIED TO DATA ON YELLOWSTONE GRIZZLY BEARS

AD SURV		SUBADULT SURV		CUB SURV		REPRO
DEATHS	DAYS	DEATHS	DAYS	SURVIVE	DEATHS	RATE
1	521	1	488	2	0	0.5
0	730	1	363	2	0	0.6
0	863	0	130	2	0	0.888889
1	143	0	259	3	0	1
1	126	0	342	2	0	1
0	732	0	187	2	1	0.571429
1	447	0	930	1	1	0.714286
0	1684	0	190	3	0	0.571429
0	2512	1	248	2	0	0.833333
0	1100	1	209	3	0	0.8
0	540	0	238	2	0	0.666667
0	977	0	967	1	2	1
0	1056	0	49	3	0	0.833333
0	2182	1	80	2	0	0.571429
0	1462	0	465	2	0	0.6
0	1097	0	229	3	0	0.777778
0	17	0	259	2	0	1
0	1053	0	952	2	0	0.5
0	961	0	951	1	1	1
0	713	0	441	0	2	0.666667
0	715	0	196	2	0	0.666667
0	537	0	89	3	0	1.2
0	775	0	70	1	2	0.5
0	924	5	8332	1	0	0.4
0	951			1	0	0.5
0	699	surv	0.803241	1	1	0.4
0	831	BOOT		3	0	1.25
0	840	MEAN	0.798568	2	0	0.666667
0	29			3	0	0.75
0	868			1	1	1
0	647			58	11	1
0	650					0.333333
0	655			SURVIVAL	0.8405797	0.333333
0	935			BOOT		1
0	491			MEAN	0.842607	0.75
1	58					0.666667
0	786					0.333333
0	500					0
0	415					1.5
5	31222					0
						0.333333
S=	0.943219					0
BOOT						0
MEAN	0.942002					0.5
						1
						0.666667
						1.5
						1
					AVE	0.694692
				BOOT	MEAN	0.694636

As noted previously, cub survival was multiplied by subadult survival to get the data to enter for JSURV in the final program, using the equation:

$$\text{JSURV} = (\text{cub surv})(\text{subadult surv})^3$$

The final program was adapted from APPLMB by changing the formula attached to RES, so that the reproductive rate was divided by 2, to use female births per female. Also, there is a single entry for SJUV due to the calculation shown above being done on a separate spreadsheet. One has to be sure to get the proper equation in place for RES, but his makes it possible to use one program here instead of several variants. The first part of the final program follows, showing the input data and estimates of lambda (the full output contains 1,000 values).

RES = =G5^E5-B5*G5^(E5-1)-(D5/2)*C5*(1-(B5/G5)^(F5-E5+1))

APPLMBbear

	A	B	C	D	E	F	G	H	I	J	K
1											
2		NOBS									
3		1000							Estimate		No. of
4		S	SJUV	m	a	w	lbd	RES	of lambda	Residuals	iterations
5		0.966442	0.29994	0.689666	4	20	1.[illegible]86	9.52E-05			
6	1	0.94974	0.353568	0.761789					1.044361	-9.21E-05	73
7	2	0.91558	0.36263	0.63991					1.006698	5.96E-05	110
8	3	0.949687	0.36766	0.61701					1.026748	9.6E-05	103
9	4	0.972916	0.404718	0.753158					1.072843	-9.71E-05	131
10	5	0.956332	0.584835	0.661657					1.089553	-9.25E-05	110
11	6	0.954833	0.403478	0.679315					1.049542	-9.49E-05	123
12	7	0.920598	0.555587	0.654663					1.059422	-8.87E-05	85
13	8	0.92984	0.529066	0.642824					1.057422	-9.09E-05	92
14	9	0.969526	0.373506	0.750248					1.061492	-9.52E-05	139
15	10	0.946021	0.378807	0.68292					1.037859	9.57E-05	49
16	11	0.967661	0.361823	0.636359					1.04	4.53E-05	1
17	12	0.978505	0.553779	0.679117					1.100384	-8.95E-05	131
18	13	0.942578	0.362558	0.743667					1.03999	9.01E-05	2
19	14	0.974791	0.470335	0.799074					1.097914	-9.37E-05	127
20	15	0.966787	0.479183	0.693304					1.078493	-9.13E-05	120

Observant readers will note that this example varies a little from that shown earlier in that the box for RATE was inserted after this example was run (the starting value was originally in the program, and has now been moved to the spreadsheet.

The bootstrap value of lambda was 1.053, as is the value calculated from the original data, using SOLVER. The 95% confidence interval for lambda is 0.97 to 1.12. These wide limits are a consequence of the limited data on subadult survival. Using the delta method shows that 77% of the variability of the estimate of lambda is due to the small sample for subadult survival (Eberhardt 2002:Table 2). Radiotelemetry has been continued on the population since the above data were collected, and hopefully a better estimate will eventually be available (a trend estimate provided by Eberhardt et al. 1999) gives a somewhat better notion of the likely status of the population). It is assumed here that the different sets of data used for inputs are independent. If they are not independent, the confidence limits will be narrower than they should be. Hence, it is a good idea to check the correlations between the bootstraps. In this example, they were all quite small (less than 0.10).

Survival estimation with the Jolly-Seber method

Estimating survival is an essential feature of any study of population dynamics. The program used here (BOOTSURV) follows the approach of Seber (1982), and provides survival estimates by the Jolly-Seber method. There is a necessary dichotomy in applications of capture-recapture methods. The original theory assumes that all previously unmarked animals are marked and released as they are captured. This is usually not feasible with large mammals, where studies usually begin with a substantial marked sample, with new "batches" marked when it is practicable to do so. When the actual capture-recapture study begins, the first capture of a marked individual is then effectively regarded as the initial capture and marking.

The essential equation for the Jolly-Seber method is (Seber 1982:200):

$$\hat{M}_i = \frac{R_i z_i}{r_i} + m_i \qquad (i = 2,3,...,s-1)$$

Here R_I denotes the number of marked individuals released at time i, z_i is the number captured before and after i, but not at time i, r_i denotes those animals released at i, and caught again later, and m_i are previously marked animals caught at time i. The z_i and r_i are usually obtained by constructing two intermediate tables as illustrated by Seber (1982:206). A different approach is used in the present program, in that z is estimated from two other parameters, i.e., z = C – k in the table produced by BOOTSURV.

The basic data are represented by a rectangular matrix in which the rows represent the capture history of individuals with 1 denoting capture on the ith occasion and zero, no capture, and the columns contain the data for the capture occasions.

The Jolly-Seber method estimates survival as (Seber 1982, eq.(5.9), p. 200):

$$\hat{\phi}_i = \frac{\hat{M}_{i+1}}{\hat{M}_i - m_i + R_i} \qquad (i = 2,3,...s-2)$$

BOOTSURV uses 3 different worksheets. The first one (DATA) contains the original data, which will be, as noted above, a matrix of zeros and ones. The data must start in the upper left-hand corner of the sheet, and no other information appears on the sheet – it is called by the programs associated with the other two sheets. The second sheet is titled “Original Data” and calculates survival by the equations given above. The box titled “NROWS” contains the number of rows of data (number of individuals for which data is available), and the second box, titled “NCOLS” contains the number of columns of data, which correspond to the number of capture occasions. When you run the program associated with this sheet, it prints out the data on the right, permitting one to check to be sure that the data are correct. The program is illustrated with data on the Florida manatee collected at the Blue Springs site. There were 15 capture occasions, and the equation for survival above produces survival estimates for occasions 2,…,s-2 (where s is the number of capture occasions), and these 13 values appear in the right-hand column of output, and are averaged below. You start the program either by using the Macro call in the Tools menu, or by hitting Alt & F8. When you do this a box comes up that says “ Bootstrap1” and “Original data”. Be sure to highlight “Original data” in order to carry out the analysis described above. “Bootstrap1” is automatically highlighted, so if you just hit “return” you will get an error message (just press “End” in the box that comes up to get back to the worksheet). You must highlight the appropriate name to run the corresponding program. The picture on the next page shows the Original Data worksheet for the Blue Springs manatee data.

Once you have run the analysis on the sheet titled “Original data”, switch to the sheet titled “Bootstraps” (a picture of that sheet follows the Original Data sheet below). The box labeled “NBOOT” in the upper left-hand corner determines the number of bootstraps. I suggest starting out with 30 until you are sure you have the program running properly. The program obtains the number of rows and the number of columns from the previous sheet (“Original data”) and operates only on that part of the data. However, if it has been used for a different data set at some previous time, there may be more rows of data than called for by “Original data”, which can be confusing. I suggest clearing the columns below C,k,R,m and r before you start. If you clear more columns you will destroy the outputs you want (just close down and start over to restore the computations). When you run the sheet the above columns are filled in and the survival rate is computed under “surv”. This program is set up to tally only an average survival rate, so you need to be sure that an average of the presently used data is calculated below the “surv” column and that the box under “PARAM INPUT” is set equal to this average. Another program is available that prints out ALL of the bootstrap survival estimates and does not average them. It can be used to study the pattern of survival over time, but is not considered here for the present. The data shown here are in the copy of the program included with this Appendix, so readers can experiment with running the program with this data.

surv boot bs.xls

	A	B	C	D	E	F	G	H	I	J	K	L
1	NROWS	**C**	**k**	**R**	**m**	**r**	**M-hat**	**p-hat**	**JSsurv**			
2	68	13	13	16	13	16	13.00	1				
3	NCOLS	16	16	19	16	19	16.00	1	1			
4	15	19	18	21	18	21	19.00	0.94737	1			
5		21	20	23	21	21	22.10	0.95043	1.0043			
6		21	21	26	22	25	22.00	1	0.913			
7		25	18	20	18	20	25.00	0.72	0.9615			
8		25	23	32	25	29	27.21	0.91888	1.0077			
9		31	30	36	30	35	31.03	0.96685	0.9071			
10		34	34	40	36	36	36.00	1	0.9722			
11		35	34	39	35	38	36.03	0.97151	0.9007			
12		38	37	40	38	39	39.03	0.97372	0.975			
13		38	35	42	37	40	40.15	0.92154	0.9787			
14		40	39	50	42	45	43.11	0.97423	0.9548			
15												
16												
17								AVERAGE	0.9646			
18												
19												

OriginalData | BOOTSTRAPS

surv boot bs.xls

	A	B	C	D	E	F	G	H	I	J	K
1	NBOOT	66	**C**	**k**	**R**	**m**	**r**	surv	M-hat	0.96465	PARAM
2	1000	51	10	10	13	10	13		10	0.95478	INPUT
3		36	13	13	18	13	18	1	13	0.96482	0.9651
4		12	18	17	19	17	19	1	18	0.96556	
5		55	19	18	20	19	19	1.003	20.05	0.96899	
6		36	19	19	21	20	20	0.95	20	0.98281	
7		48	20	14	15	14	15	0.952	20	0.96331	
8		22	20	20	33	21	30	1	21	0.97514	
9		62	30	30	37	30	34	0.909	30	0.96799	
10		51	34	34	39	34	36	0.919	34	0.96325	
11		6	35	28	33	29	32	0.929	36.22	0.97164	
12		49	37	37	40	39	38	0.97	39	0.97102	
13		46	38	36	40	36	40	0.95	38	0.96454	
14		53	32	32	52	42	38	1	42	0.95811	
15		60								0.97328	
16		19					AVE	0.965		0.98007	

OriginalData | **BOOTSTRAPS**

Another example of the use of the approximation to the Lotka-Leslie equation

The Blue Springs manatee data provide another example of the use of APPLMB, using the survival data obtained above with the Jolly-Seber method. The bootstrap results were used to introduce use of program APPLMB and the survival data were used to illustrate the use of BOOTSURV above, with the original survival data incorporated in the appended copy of BOOTSURV. The bootstrap means were very close to the estimates from original data, with the adult survival estimate from original data being 0.965, the same as the bootstrap mean. The same applies to reproductive data with both averages being 0.151. The original data for first year survival gave 0.822, while the bootstrap mean was 0.821. The survival data appear in the picture of the Original Data sheet above, while first year survival was computed from 37 survivals from 45 individuals tagged. The first year survival was bootstrapped by using an input table with 37 values of 1, and 8 zeros. The reproductive rate data (births per female per year) are shown below for use by anyone wanting to duplicate the bootstrapping.

0.3333
0.3571
0.25
0.5
0.3077
0.3333
0.2
0.2222
0.1667
0.2857
0.2857
0.3333
0.3333
0.5
0.3333
0
0.25
0.3333
0.2
0.5

Estimating a parameter using values of lambda from trend data.

When data are not available to estimate a parameter other than lambda (early survival, for instance), it may be possible to substitute estimates of lambda from trend data. This is readily done by rearranging eq.(11.9).

$$\lambda^{a} - s\lambda^{a-1} - l_a m \left[1 - \left(\frac{s}{\lambda}\right)^{w-a+1}\right] = 0 \qquad (11.9)$$

For the Blue Springs manatee data, l_a was estimated by assuming survival from the first year to the age of first reproduction (a) was at the adult rate, so that

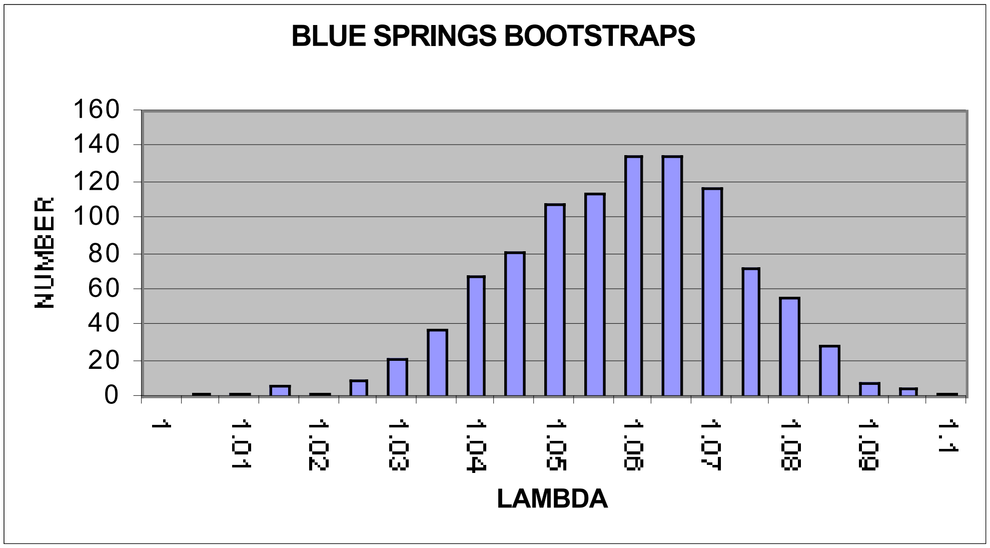
$$l_a = s_{juv} s^{a-1}$$

Thus we can estimate first year survival from:

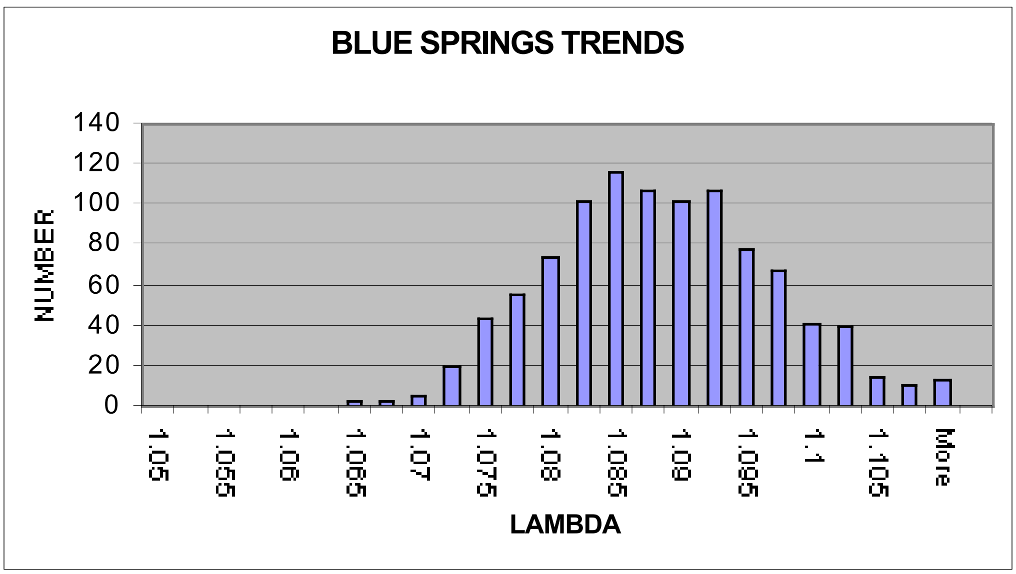
$$s_{juv} = \frac{\lambda^a - s\lambda^{a-1}}{s^{a-1} m[1 - (\frac{s}{\lambda})^{w-a+1}]}$$

and m can be estimated by a similar rearrangement. Estimating adult survival (s) would require an iterative solution, but, while APPLMB can be revised to do that, I do not recommend it because adult female survival is such an important parameter that it should be estimated directly. Bootstrapping can be arranged by using the above equation. Nineteen years of trend data were available for Blue Springs manatees and were bootstrapped with BOOT2 to get 1,000 estimates of λ from trend data. These can then be used with the available bootstraps of adult female survival and reproductive rate, arranged on a worksheet, and first year survival estimated from the above equation.

Most of the estimates of SJUV are unreasonably high. Why this should be the case can be appreciated by comparing a frequency plot of the bootstrap estimates of lambda estimated from reproductive and survival data with one estimated by bootstrapping trend data from Blue Springs manatees. These plots are shown below. The trend data give much higher estimates of λ than were obtained by bootstrapping reproductive and survival data. Eberhardt and O'Shea (1995) reported indications of immigration to the two sites for which trend data were obtained (Crystal River and Blue Springs). Estimates of early survival from the Crystal River site were not available, so the main indication that there might be immigration likely has to come from the failure of the above program to produce satisfactory estimates for that site, too, along with the impressions recorded by Eberhardt and O'Shea (1995).



Frequency plot of bootstraps of estimates of λ based on reproductive and survival data.



Frequency plot of estimates of λ obtained by bootstrapping trend data.

Bootstrapping multiple regressions

Multiple regressions may be useful in assessing trend data, as indicated in Section 9.10. Hence it is worthwhile to have two programs to bootstrap multiple regressions. MREGNONP takes the usual bootstrap approach, in which the set of observations is randomly sampled n times with replacement, and a multiple regression is fitted to each of the B bootstrap samples. The calculations are done on the sheet titled “DATA”, and the bootstrapped values of the regression calculations appear on a sheet titled “BOOT”. It is convenient to attach another sheet containing the usual EXCEL regression calculations, and that sheet is labeled “REGRDATA” in the attached programs, which are based on the grizzly bear data of Example 9.6. The regressions are calculated by matrix operations which appear of the right of the sheet titled “DATA”. Up to 5 independent variables can be accomodated, but note that the box under NVAR should show p+1 where p is the number of independent variables. The first column (labeled Xo) has to contain a column of ones for calculating the intercept by matrix methods. The sheet labeled “DATA” is as follows (without the matrix calculations which are off to the right of the sheet):

MREGNONP.xls

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	NOBS	NBOOT	NVAR												
2	20	20	4					RAND							
3	Y	X0	X1	X2	X3	X4	X5	NO.	Y	X0	X1	X2	X3	X4	X5
4	2.8332	1	1	1.6	7.6			15	2.944	1	13	2.12	-15.7		
5	2.5649	1	2	1.5	10.4			6	2.197	1	3	1.3	-18.4		
6	2.1972	1	3	1.3	-18.4			15	2.944	1	13	2.12	-15.7		
7	2.5649	1	4	1.1	3.9			4	2.833	1	1	1.6	7.6		
8	2.4849	1	5	1.4	1.6			20	2.996	1	18	1.67	-6.33		
9	2.6391	1	6	1.6	-4.4			21	2.996	1	19	1.47	3.27		
10	2.3979	1	7	1.6	-9.8			19	3.136	1	17	1.65	0.47		
11	2.5649	1	8	1.2	9.6			8	2.485	1	5	1.4	1.6		
12	2.8332	1	9	2.3	2.9			7	2.565	1	4	1.1	3.9		
13	3.2189	1	11	3.12	-0.6			17	3.219	1	15	1.95	8.97		
14	2.565	1	12	1.64	3.5682			9	2.639	1	6	1.6	-4.4		
15	2.944	1	13	2.12	-15.732			10	2.398	1	7	1.6	-9.8		
16	2.773	1	14	1.86	-3.1318			9	2.639	1	6	1.6	-4.4		
17	3.219	1	15	1.95	8.9682			11	2.565	1	8	1.2	9.6		
18	3.178	1	16	2.65	-4.7318			4	2.833	1	1	1.6	7.6		
19	3.136	1	17	1.65	0.4682			11	2.565	1	8	1.2	9.6		
20	2.996	1	18	1.67	-6.3318			8	2.485	1	5	1.4	1.6		
21	2.996	1	19	1.47	3.2682			14	2.565	1	12	1.64	3.57		
22	2.833	1	20	1.47	-5.3318			12	2.833	1	9	2.3	2.9		
23	3.497	1	21	1.96	12.668			13	3.219	1	11	3.12	-0.6		
24															
25															

The second program, MREGRPARA, uses the parametric approach to the same data set, with the deviations from regression and “predicted” values calculated by the usual EXCEL regression program (attached as sheet “REGRDATA”) in the lefthand 2 columns. The deviations are bootstrapped and attached to the predicted values to generate a new y-value next to the values of the independent variables as follows:

MREGPARA.xls

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	NOBS	NBOOT	NVAR												
2	20	20	4					RAND							
3	DEV.	PREDICT	DEV SELECTED	ORIG Y				NO.	Y	X0	X1	X2	X3	X4	X5
4	0.243	2.590674	-0.013619556	2.8332				18	2.577	1	1	1.6	7.6		
5	-0.05	2.614539	-0.093717318	2.5649				22	2.521	1	2	1.5	10.4		
6	0.02	2.177489	0.091441279	2.1972				17	2.269	1	3	1.3	-18.4		
7	0.114	2.450921	-0.019793207	2.5649				21	2.431	1	4	1.1	3.9		
8	-0.066	2.551125	0.135475412	2.4849				23	2.687	1	5	1.4	1.6		
9	0.083	2.555726	-0.122377096	2.6391				11	2.433	1	6	1.6	-4.4		
10	-0.126	2.523973	-0.093717318	2.3979				22	2.43	1	7	1.6	-9.8		
11	-0.122	2.687326	0.242539169	2.5649				4	2.93	1	8	1.2	9.6		
12	-0.139	2.972321	0.162972299	2.8332				15	3.135	1	9	2.3	2.9		
13	-0.03	3.248444	0.135475412	3.2189				23	3.384	1	11	3.12	-0.6		
14	-0.3	2.86451	0.083331347	2.5649				9	2.948	1	12	1.64	3.568		
15	0.163	2.781467	0.078645218	2.9444				20	2.86	1	13	2.12	-15.73		
16	-0.13	2.902119	0.162972299	2.7726				15	3.065	1	14	1.86	-3.132		
17	0.091	3.127435	0.160989036	3.2189				19	3.288	1	15	1.95	8.968		
18	-0.014	3.191673	0.078645218	3.1781				20	3.27	1	16	2.65	-4.732		
19	0.161	2.974505	0.162972299	3.1355				15	3.137	1	17	1.65	0.468		
20	0.079	2.917087	-0.093717318	2.9957				22	2.823	1	18	1.67	-6.332		
21	-0.02	3.015525	0.019735874	2.9957				6	3.035	1	19	1.47	3.268		
22	-0.094	2.926931	-0.066218453	2.8332				8	2.861	1	20	1.47	-5.332		
23	0.135	3.361032	-0.139107279	3.4965				12	3.222	1	21	1.96	12.67		

Multiple regression calculations are then done on this data set.