Taking the Next Step:  
Replicating Eff’s Autocorrelation Study to Test for Female Subsistence Contributions

Sarah I. Baitzel

2007, UCI COURSE, Dr. D. White

One of the most well known characteristics of cultural anthropology is the use of qualitative data to describe, analyze and interpret the societies, which they study. Particularly with the rise of Geertzian “thick description” (1973), the field of cultural anthropology has become increasingly reliant on textual evidence. Quantitative data and their interpretation are less frequent and may perhaps be easily overlooked by followers of the descriptive type with the argument that quantitative data and statistical analysis lack an emic perspective and therefore the ability to really understand the underlying working mechanisms of a given society. Such judgment, however, arises only from the misconceptions about the use of statistics in anthropology. Like any other field in the social sciences, anthropology has seen its fair share of faulty statistical application and analysis. Most commonly, scholars lack accurate and precise working hypotheses, so their tests and results remain without value to their research question. However, statistical analysis, rather than looking for contextual connections between entities, has the ability to take change perspectives, examining the data from a purely mathematical angle. This produces results that may not be visible or apparent to the research from a textual descriptive approach, and therefore opens doors for new enquiries and studies. In this paper, I will replicate the statistical analysis by E. Anthon Eff on autocorrelation of Variable 826 from Murdock and White’s Ethnographic Atlas (Average Adult Female Contribution to Subsistence) in R. I will then extend the original study to test for a better
correlation of variables; by carefully choosing the variables, I will increase the degree to which the variation in the other cultural variables accounts for the variation in V826.

Especially in the field of cross-cultural studies, quantitative methods are of central concern to the work of cultural anthropologists. Textual evidence most times cannot be applied across cultural borders, while quantitative data serves this purpose particularly well. By comparing societies from the past and the present, scholars seek to discover underlying human universals and “natural laws” of culture (Divale 1984:219). The problem for cultural anthropologists is that different cultures and societies do not exist in a vacuum. They are not separate Petri dishes in a laboratory that never come into contact with one another. Instead, we may assume that human societies are linked through one or more of the following processes: migration, diffusion and common descent (Eff 2004:153). This was Galton’s counter-argument to Tylor’s cross-cultural study with which he intended to prove that a number of cultures could be placed in an evolutionary sequence based on the available information about marriage and descent. A correlation of variables due to diffusion and borrowing would imply that the disturbance terms in the applied linear regression analysis are not completely independent of each other. This causes “the estimated standard errors of the coefficients to be biased” (Eff 2004:153).

To avoid Galton’s problem, Eff employs Moran’s I statistic, borrowed from spatial econometrics. This allows him to test for phylogenetic relatedness (common descent) and spatial proximity (distance). Eff is not the first author attempting to solve Galton’s problem. A study by H. Carroll (1974) investigated whether the connection between a society’s dependence on agriculture and the presence of patrilocal, virilocal or avuncular residence patterns was in fact functional or whether it was due to diffusion.
Carroll used a linked pair test and a double language boundary test to test for diffusion of traits among cultures. His results from the linked pair test indicated, “the variables are functionally associated and there is a tendency for both to diffuse together” (1974:185). The double language boundary test accounts for cultural diffusion and allows for the selection of culture areas that are not significantly affected by borrowing. This test confirmed that semi-diffusional and ideographic-diffusional relationship of the variables.

The objective of finding universal human laws through cross-cultural research seeks to identify not diffusional relationships. Instead, the goal is to find functional relationships, which can be identified on the basis that they are found on a worldwide basis in a statistically significant association (Greenbaum 1974:151). Although random and stratified sampling strategies can to some extend alleviate the problem of autocorrelation based on diffusion and migration, spatial proximity and linguistic relatedness must still be tested for to lessen the impact of Galton’s problem on cross-cultural studies.

Unlike Carroll, we today have the advantage of advanced statistical computer programs that supply us with a number of tools for running different tests. To replicate Eff’s study, I will use R, in particular the “spdep” package (“Spatial dependence: weighting schemes, statistics and models”). This package contains – among other things – a tool for analyzing spatial lag in ordinary least square regression models. In order to test for diffusion, Eff uses the location (based on latitude and longitude), the language families, and the levels of cultural complexity of 186 cultures as indicated in the Ethnographic Atlas (Murdock and White 1969) (Eff 2004:154). In R, the spatial and linguistic information is converted into weight matrices (langmat.dbf and distmat.dbf),
which are then converted into one matrix. This is an excerpt from the matrix showing the first ten variables:

\[
\text{lds}[1:10, 1:1] =
\begin{bmatrix}
X10 & X9 & X8 & X7 & X6 & X5 & X4 & X3 & X2 & X1 \\
[1,] 0.699 & 0.691 & 0.735 & 0.761 & 0.758 & 0.819 & 0.842 & 0.834 & 0.894 & 0.000 \\
[2,] 0.762 & 0.759 & 0.805 & 0.838 & 0.831 & 0.892 & 0.938 & 0.849 & 0.000 & 0.894 \\
[3,] 0.775 & 0.752 & 0.807 & 0.817 & 0.735 & 0.771 & 0.844 & 0.000 & 0.849 & 0.834 \\
[4,] 0.805 & 0.803 & 0.852 & 0.889 & 0.860 & 0.901 & 0.000 & 0.844 & 0.938 & 0.842 \\
[5,] 0.758 & 0.769 & 0.797 & 0.833 & 0.917 & 0.000 & 0.901 & 0.771 & 0.892 & 0.819 \\
[6,] 0.776 & 0.799 & 0.809 & 0.841 & 0.000 & 0.917 & 0.860 & 0.735 & 0.831 & 0.758 \\
[7,] 0.898 & 0.896 & 0.952 & 0.000 & 0.841 & 0.833 & 0.889 & 0.817 & 0.838 & 0.761 \\
[8,] 0.941 & 0.924 & 0.000 & 0.952 & 0.809 & 0.797 & 0.852 & 0.807 & 0.805 & 0.735 \\
[9,] 0.948 & 0.000 & 0.924 & 0.896 & 0.799 & 0.769 & 0.803 & 0.752 & 0.759 & 0.691 \\
[10,] 0.000 & 0.948 & 0.941 & 0.898 & 0.776 & 0.758 & 0.805 & 0.775 & 0.762 & 0.699 \\
\end{bmatrix}
\]

The next step is the creation of a data frame with the variables, which Eff thought to be related (either functionally or by borrowing) to female contributions to subsistence. The variable for average female contributions to subsistence is a percentage of weighted sums of variables combining the importance of a certain subsistence strategy in society (in percentage) with the percentage of female contribution to subsistence (SCCS). It is not a “measure of the total amount of work done by women in comparison to men in a society”, but rather, it measures the “relative importance females play in harnessing food energy in a society” (Divale 1984:148).

Eff chooses the following variables to account for variation in female contributions to subsistence: the importance of fishing and hunting (in percent as they contribute to overall subsistence), total pathogen stress, niche rainfall, polygamy, sex of parent in residence: early boy, fixity of residence, land transport, political integration, and social stratification. Although Eff does not explain the reasons why he chose these particular variables, they do represent most aspects of society in terms of subsistence, residence, socio-political structure, health, and environment. The following table
summarizes the numeric values for each variable as they are listed in the Ethnographic
Atlas:

```
summary(ffd)
femsubs fishimp huntimp pathstress
Min. : 0.00 Min. : 0.00 Min. : 0.00 Min. : 7.00
1st Qu.:21.00 1st Qu.: 5.00 1st Qu.: 5.00 1st Qu.: 9.00
Median :32.50 Median : 5.00 Median : 5.00 Median :12.00
Mean :32.74 Mean :15.89 Mean :15.28 Mean :12.55
3rd Qu.:42.25 3rd Qu.:25.00 3rd Qu.:25.00 3rd Qu.:15.25
Max. :79.00 Max. :90.00 Max. :80.00 Max. :21.00
```
```
rainfall polygamy eboysxp fixres
Min. :1.000 Min. :1.000 Min. :2.000 Min. :1.000
1st Qu.:1.750 1st Qu.:3.000 1st Qu.:3.000 1st Qu.:2.000
Median :3.000 Median :3.000 Median :3.000 Median :5.000
Mean :3.278 Mean :3.128 Mean :3.367 Mean :3.722
3rd Qu.:5.000 3rd Qu.:4.000 3rd Qu.:3.250 3rd Qu.:5.000
Max. :7.000 Max. :4.000 Max. :5.000 Max. :5.000
```
```
landtrans polinteg socstrat
Min. :1.000 Min. :1.000 Min. :1.000
1st Qu.:1.000 1st Qu.: 2.000 1st Qu.:1.000
Median :1.000 Median :3.000 Median :2.000
Mean :1.783 Mean :2.944 Mean :2.439
3rd Qu.:2.000 3rd Qu.:4.000 3rd Qu.:4.000
Max. :5.000 Max. :5.000 Max. :5.000
```

The missing values have to be excluded from the matrix, so that only 180 cultures remain
in the sample.

The variables are now analyzed for any possible autocorrelation using the
lagsarmal tool, which examines the “maximum likelihood estimation for spatial
simultaneous autoregressive lag” (R manual 2008).

|            | Estimate  | Std. Error | z value | Pr(>|z|) |
|------------|-----------|------------|---------|---------|
| (Intercept) | 45.038893 | 15.118398  | 2.9791  | 0.0028912 |
| fishimp    | -0.198235 | 0.064857   | -3.0565 | 0.0022395 |
| huntimp    | -0.429073 | 0.079107   | -5.4239 | 5.83e-08  |
| pathstress | -1.253772 | 0.397589   | -3.1534 | 0.0016136 |
| rainfall   | -2.308783 | 0.660950   | -3.4931 | 0.0004774 |
| polygamy   | 2.920663  | 1.703061   | 1.7149  | 0.0863545 |
| eboysxp    | 5.429865  | 1.723439   | 3.1506  | 0.0016294 |
The highlighted z-scores are of not of interest, because they have values <0.05, which indicates that their correlation with the other variables and with female contribution to subsistence is relevant. The variables with values falling above 0.05 will be subject to change, because it is likely that other variables will show greater significance. The R-squared value of this combination of variables is 0.3117, meaning that 31% of the variation in the ten variables accounts for variation in variable 826. The objective of this extended study is to replace variables with weaker significance with others until the R-squared value exceeds 0.40. It should be noted here that the number of variables will not be changed, because this would affect the autocorrelation and invalidate comparative results.

Since polygamy has a relatively high z-score, I will begin by replacing this variable. Polygamy is the broad description of the practice of having multiple spouses. The two variations within polygamy are polyandry (women with several male spouses) and polygyny (men with several female spouses). When we examine polygamy (v79) we see that the majority of cultures in the SCCS are polygynous rather than polyandrous (2:163). We will therefore replace v79 with v877 for polygyny. This gives us the following results in R:

|       | Estimate | Std. Error | z value | Pr(>|z|) |
|-------|----------|------------|---------|---------|
| (Intercept) | 63.650007 | 15.996722  | 3.9789  | 6.922e-05 |
| fishimp  | -0.181659 | 0.065371   | -2.7789 | 0.0054546 |
| huntimp  | -0.405407 | 0.078914   | -5.1373 | 2.787e-07  |
| pathstress | -1.433479 | 0.420704  | -3.4073 | 0.0006560  |
| rainfall | -2.299225 | 0.680844   | -3.3770 | 0.0007328  |
| polygyny | 1.759947  | 0.675954   | 2.6036  | 0.0092237  |
| eboysxp  | 4.442431  | 1.973220   | 2.2514  | 0.0243627  |
As we can see, the change from polygamy to polygyny has changed the z-score from 0.08 to 0.009, suggesting that there is a statistically better correlation between polygyny and female subsistence contribution than with polygamy. The z-scores for the surrounding variables have also been altered by the change, and seem to have become more pronounced. Only social stratification, which scored a 0.02 value with polygamy, now only shows 0.08. The R-squared value has also changed from 0.3117 to 0.3256. This means that the variation in the variables used now account for 32% rather than 31% of variation in v826.

Next, I replace v150 (fixity of residence) with v156 (density of population). Murdock and White use fixity of residence as an indicator of cultural complexity. It may be argued that density of population is an equally useful variable for measuring cultural complexity as the agglomeration of people in one place exponentially increases the amount of interactions between individuals, ultimately leading to increasing social complexity.

|          | Estimate | Std. Error | z value | Pr(>|z|) |
|----------|----------|------------|---------|---------|
| (Intercept) | 61.573447 | 15.751933 | 3.9089  | 9.270e-05 |
| fishimp   | -0.181646 | 0.065300  | -2.7817 | 0.0054073 |
| huntimp   | -0.419998 | 0.081808  | -5.1340 | 2.837e-07 |
| pathstress| -1.440055 | 0.420787  | -3.4223 | 0.0006210 |
| rainfall  | -2.191237 | 0.663885  | -3.3006 | 0.0009647 |
| polygyny  | 1.843448  | 0.676597  | 2.7246  | 0.0064382 |
| eboysxp   | 4.154045  | 1.972950  | 2.1055  | 0.0352478 |
| denspop   | -1.650607 | 0.950872  | -1.7359 | 0.0825839 |
| landtrans | -2.404004 | 1.133387  | -2.1211 | 0.0339150 |
| polinteg  | 2.801116  | 1.379594  | 2.0304  | 0.0423167 |
| socstrat  | -2.259084 | 1.106250  | -2.0421 | 0.0411406 |
Although the switch from fixity of residence to population density did not positively affect the value for population density, it did affect all other variables in a positive manner. The R-squared value has increased insignificantly from 0.3256 to 0.3262.

We may suspect that female subsistence contributions correlate strongly with the family structure and interactions. It is therefore safe to assume that the interaction between parent (in particular the mother) and child (especially during infancy and early childhood) can reflect some of the mother’s activities. For this reason, v157, which – like v156 – is a measure of cultural complexity and therefore redundant, is replaced with v27. This variable describes bodily contact during late infancy, ranging from precautionary and occasional contact to almost constant contact.

|            | Estimate | Std. Error | z value | Pr(>|z|) |
|------------|----------|------------|---------|---------|
| (Intercept)| 54.394545| 18.338676  | 2.9661  | 0.003016|
| fishimp    | -0.181969| 0.087233   | -2.0860 | 0.036977|
| huntimp    | -0.495474| 0.107736   | -4.5990 | 4.246e-06|
| pathstress | -1.382850| 0.526253   | -2.6277 | 0.008596|
| rainfall   | -1.652613| 0.799066   | -2.0682 | 0.038623|
| polygyny   | 0.945313 | 0.824890   | 1.1460  | 0.251801|
| eboysxp    | 7.524213 | 2.548917   | 2.9519  | 0.003158|
| denspop    | -2.103719| 1.219341   | -1.7253 | 0.084475|
| landtrans  | -1.372187| 1.316436   | -1.0424 | 0.297249|
| bodcontctei| -0.682382| 1.545571   | -0.4415 | 0.658845|
| socstrat   | -2.183064| 1.247583   | -1.7498 | 0.080147|

Although the addition of v27 seems to have had a negative effect on the significance of most other variables in the study, it increased the R-squared value from 0.3262 to 0.3727. This now means that the variation of the combined number of variables in this group, although individual variables have not shown increased significance, not account for 6% more than it did in the beginning. This is quite a noteworthy change.
The last variable to replace is land transport. To maintain a variable related to residence and to replace fixity of residence, I have chosen residence pattern (v572); this variable contains two sub-categories: societies that favor the formation of fraternal interest groups (avuncular, patrilocal and virilocal societies), and those that do not (matrilocal, ambilocal and neolocal groups).

| Variable     | Estimate  | Std. Error | z value | Pr(>|z|) |
|--------------|-----------|------------|---------|---------|
| (Intercept)  | 78.30261  | 27.73950   | 2.8228  | 0.004761|
| fishimp      | -0.14894  | 0.10531    | -1.4142 | 0.157299|
| huntimp      | -0.49520  | 0.14894    | -3.3248 | 0.000885|
| pathstress   | -0.69085  | 0.73500    | -0.9399 | 0.347250|
| rainfall     | -0.36871  | 1.07358    | -0.3434 | 0.731267|
| polygyny     | 1.73732   | 1.24258    | 1.3982  | 0.162067|
| eboysxp      | 5.99237   | 3.41752    | 1.7534  | 0.079528|
| denspop      | -2.28637  | 1.51408    | -1.5101 | 0.131024|
| respatt      | 7.42570   | 4.24252    | 1.7503  | 0.080066|
| bodcontctei  | -1.30172  | 1.98536    | -0.6557 | 0.512044|
| socstrat     | -5.56993  | 1.78633    | -3.1181 | 0.001820|

With this final variable exchange, the R-squared value has now reached 0.4588. Since this exceeds the originally set goal of 0.40, I will now discontinue the variable exchange and proceed to a discussion of the results and the implications of this study for future research on cross-cultural studies.

Analyzing the estimated parameters of the variables, we can now draw certain conclusions about the characteristics of societies in which average female contribution to subsistence is higher. This apparently occurs in societies where fishing is a less important subsistence strategy, but in societies where hunting is more important, females contribute less to the overall energy harnessed. It seems self-explanatory, because women are less likely to participate in hunting activities than men. The total pathogen stress of these societies is low, an observation that may be functionally related to low rainfall levels. It is important, however, not to equate statistical correlation directly with
functional correlation. In societies with higher average female contributions to
subsistence, men tend to have multiple wives, mothers spent a large amount of time with
young boys, but there seems to be a low level of bodily contact for infants. The level of
social stratification is not high, populations are generally sparsely distributed, and group
members favor the formation of fraternal interest groups along virilocal residence
patterns.

Based on this information, it is possible to come back to the ethnographic record
and identify societies that follow this particular pattern. Low rainfall levels indicate a
semi-arid to arid environment, possibly located in the subtropical or continental
temperate zones. Assuming that hunting and fishing are less important to these societies,
we may suspect that they are primarily relying on either gathering or agriculture. The
family structure is male-centered and women spend a lot of time with young male
children, possibly reflecting the importance of men in society in general.

The application of a lagged regression models that considers proximity and
phylogenetic relationship (using language as a proxy) enables social scientists to
eliminate one of the most important problems in cross-cultural research. While previous
studies had to rely on sampling from the Ethnographic Atlas, a process that is inherently
biased, the use of advanced statistical computer programs such as R allows researchers to
expand their sample size and to use complex analytical tools. The example shown in this
paper demonstrates how the careful selection and supplementation of variables can
incrementally increase the degree to which the variation in the chosen variables account
for the variation in the variable that is being investigated. Such a study complements
textual research in finding underlying statistical correlations.
APPENDIX: Calculations in R

R version 2.6.2 (2008-02-08)
Copyright (C) 2008 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and 'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or 'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace restored from /Users/sarahbaitzel/.RData]

> options(echo=TRUE)
> library(foreign)
> library(spdep)
Loading required package: tripack
Loading required package: sp
Loading required package: maptools
Loading required package: boot
Loading required package: spam
Package 'spam' is loaded. Version 0.13-3 (2008-04-21).

Type demo( spam) for some demos, help( Spam) for an overview of this package.

Loading required package: Matrix
Loading required package: lattice

Attaching package: 'lattice'
  The following object(s) are masked from package:boot :
    melanoma
Attaching package: 'Matrix'
  The following object(s) are masked from package:spam :
    norm
  The following object(s) are masked from package:stats :
    xtabs

Warning messages:
1: In namespaceImportFrom(self, asNamespace(ns)) :
  replacing previous import: as.matrix
2: In namespaceImportFrom(self, asNamespace(ns)) :
replacing previous import: chol
3: In namespaceImportFrom(self, asNamespace(ns)) :
replacing previous import: determinant
4: In namespaceImportFrom(self, asNamespace(ns)) :
replacing previous import: diag
5: In namespaceImportFrom(self, asNamespace(ns)) :
replacing previous import: image
6: In namespaceImportFrom(self, asNamespace(ns)) :
replacing previous import: norm
7: In namespaceImportFrom(self, asNamespace(ns)) :
replacing previous import: print
8: In namespaceImportFrom(self, asNamespace(ns)) :
replacing previous import: solve
9: In namespaceImportFrom(self, asNamespace(ns)) :
replacing previous import: summary
10: In namespaceImportFrom(self, asNamespace(ns)) :
replacing previous import: t
> lds<-read.dbf("langmat.dbf")
> lds<-read.dbf("distmat.dbf")
> lds<-as.matrix(lds)
> gg<-read.dta("SCCS_stata.dta")
> df<-
data.frame(femsubs=gg$v826,fishimp=gg$v816,huntimp=gg$v817,pathstress=gg$v1260,rainfall=gg$v855,polygyny=gg$v877,eboysxp=gg$v353,fixres=gg$v150,landtrans=gg$v154,polinteg=gg$v157,socstrat=gg$v158)
>
> kk<-as.matrix(df)
> oo<-(1,length(kk[1,]),1)
> tm<-kk%%oo
> rr<-which(tm[,1]!="NA")
> wmat<-mat2listw(lds[rr,rr])
> ffd<-df[rr,]
> summary(ffd)
defs
  femsubs fishimp huntimp pathstress rainfall
  Min. : 0.00 Min. : 0.00 Min. :7.00 Min. :1.00
  1st Qu.:21.00 1st Qu.:5.00 1st Qu.:9.00 1st Qu.:2.00
  Median :32.00 Median :5.00 Median :12.00 Median :3.00
  Mean :32.81 Mean :15.91 Mean :15.51 Mean :12.55
  3rd Qu.:43.00 3rd Qu.:25.00 3rd Qu.:15.50 3rd Qu.:5.00
  Max. :79.00 Max. :90.00 Max. :21.00 Max. :7.00
polygyny eboysxp fixres landtrans polinteg
  Min. :0.00 Min. :2.00 Min. :1.00 Min. :1.00
  1st Qu.:0.00 1st Qu.:3.00 1st Qu.:1.00 1st Qu.:1.00
  Median :1.00 Median :5.00 Median :1.00 Median :3.00
  Mean :1.829 Mean :3.354 Mean :3.697 Mean :1.777
  3rd Qu.:3.00 3rd Qu.:5.00 3rd Qu.:2.00 3rd Qu.:4.00
  Max. :6.000 Max. :5.000 Max. :5.000 Max. :5.000
socstrat
  Min. :1.00
  1st Qu.:1.00
  Median :2.00
Mean : 2.417
3rd Qu.: 4.000
Max. : 5.000
> library(spdep)
> col.lm <- lagsarlm(femsubs ~ fishimp + huntimp + pathstress + rainfall + polygyny + eboysxp + fixres + landtrans + polinteg + socstrat, data = ffd, wmat, quiet = FALSE)

Spatial lag model
Jacobian calculated using neighbourhood matrix eigenvalues
Computing eigenvalues ...

(eigen) rho:  -0.6241835  function value:  -1146.746
(eigen) rho:  -0.3808565  function value:  -1005.147
(eigen) rho:  -0.2304722  function value:  -899.543
(eigen) rho:  -0.1375295  function value:  -816.0097
(eigen) rho:  -0.0800878  function value:  -757.0601
(eigen) rho:  -0.04458688  function value:  -724.4869
(eigen) rho:  -0.02264609  function value:  -711.0509
(eigen) rho:  -0.009085945  function value:  -707.024
(eigen) rho:  0.0008470436  function value:  -706.6442
(eigen) rho:  -0.002383861  function value:  -706.5042
(eigen) rho:  -0.002548879  function value:  -706.5042
(eigen) rho:  -0.00246882  function value:  -706.5041
(eigen) rho:  -0.002468893  function value:  -706.5041
(eigen) rho:  -0.002468867  function value:  -706.5041
(eigen) rho:  -0.00246862  function value:  -706.5041
(eigen) rho:  -0.00246872  function value:  -706.5041
(eigen) rho:  -0.00246867  function value:  -706.5041
> summary(col.lm)

Call: lagsarlm(formula = femsubs ~ fishimp + huntimp + pathstress + rainfall + polygyny + eboysxp + fixres + landtrans + polinteg + socstrat, data = ffd, listw = wmat, quiet = FALSE)

Residuals:
   Min     1Q Median     3Q    Max
-31.60019  -9.70375   0.85626   9.12570  40.51744

Type: lag
Coefficients: (asymptotic standard errors)

                     Estimate Std. Error z value Pr(>|z|)
(Intercept)       63.650007  15.996722   3.9789 6.922e-05
fishimp         -0.181659   0.065371  -2.7789 0.0054546
huntimp         -0.405407   0.078914  -5.1373 2.787e-07
pathstress      -1.433479   0.420704  -3.4073 0.0006560
rainfall         -2.299225   0.680844  -3.3770 0.0007328
polygyny         1.759947   0.675954   2.6036 0.0092237
eboysxp         -4.442431   1.973220  -2.2514 0.0243627
fixres          -1.502225   0.888379  -1.6910 0.0908421
landtrans       -2.538122   1.134860  -2.2365 0.0253186
polinteg  2.329974  1.320361  1.7646 0.0776226
socstrat   -1.963763  1.129639 -1.7384 0.0821406

Rho: -0.0024689 LR test value: 0.15405 p-value: 0.6947
Asymptotic standard error: 0.0058757 z-value: -0.42018 p-value: 0.67435
Wald statistic: 0.17655 p-value: 0.67435

Log likelihood: -706.5041 for lag model
ML residual variance (sigma squared): 187.96, (sigma: 13.71)
Number of observations: 175
Number of parameters estimated: 13
AIC: 1439, (AIC for lm: 1437.2)
LM test for residual autocorrelation
test value: 0.5125 p-value: 0.47406

> pseudoR2<-cor(col.lm$fitted.values,ffd$femsubs)^2
> summary(pseudoR2)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.3256  0.3256  0.3256  0.3256  0.3256  0.3256
> df<-
data.frame(femsubs=gg$v826,fishimp=gg$v816,huntimp=gg$v817,pathstress=gg$v1260,rainfall=gg$v855,polygyny=gg$v877,eboysxp=gg$v353,respatt=gg$v572,landtrans=gg$v154,polinteg=gg$v157,socstrat=gg$v158)
> kk<-as.matrix(df)
> oo<-matrix(1,length(kk[,1]),1)
> tm<-kk%*%oo
> rr<-which(tm[,1]!="NA")
> wmat<-mat2listw(lds[rr,rr])
> ffd<-df[rr,]
> library(spdep)
> col.lm<-
lagsarlm(femsubs~fishimp+huntimp+pathstress+rainfall+polygyny+eboysxp+respatt+landtrans+polinteg+socstrat, data=ffd,wmat,quiet=FALSE)

Spatial lag model
Jacobian calculated using neighbourhood matrix eigenvalues
Computing eigenvalues ...

(eigen) rho:  -0.6242433  function value:  -520.7741
(eigen) rho:  -0.3760088  function value:  -454.1801
(eigen) rho:  -0.2225915  function value:  -408.3840
(eigen) rho:  -0.1277744  function value:  -379.2694
(eigen) rho:  -0.06917414 function value:  -364.8935
(eigen) rho:  -0.03295722 function value:  -359.6802
(eigen) rho:  -0.01057393 function value:  -358.4956
(eigen) rho:  -0.00473001 function value:  -358.4892
(eigen) rho:  -0.007353066 function value:  -358.4752
(eigen) rho:  -0.007358264 function value:  -358.4752
(eigen) rho:  -0.007358269 function value:  -358.4752
(eigen) rho:  -0.007358269 function value:  -358.4752
(eigen) rho:  -0.007358269 function value:  -358.4752
(eigen) rho:  -0.007358259  function value:  -358.4752
(eigen) rho:  -0.007358264  function value:  -358.4752
> summary(col.lm)

Call: lagsarlm(formula = femsubs ~ fishimp + huntimp + pathstress +
rainfall + polygyny + eboysxp + respatt + landtrans + polinteg +
socstrat, data = ffd, listw = wmat, quiet = FALSE)

Residuals:
       Min        1Q       Median        3Q       Max
-29.3116  -9.3692   -1.1300    8.2637   47.0522

Type: lag
Coefficients: (asymptotic standard errors)
         Estimate Std. Error z value Pr(>|z|)
(Intercept) 59.916471  20.159226  2.9722 0.0029571
fishimp     -0.116003   0.087511 -1.3256 0.1849798
huntimp     -0.420274   0.110549 -3.8017 0.0001437
pathstress  -0.788804   0.619485 -1.2733 0.2029039
rainfall    -0.871294   0.905745 -0.9620 0.3360674
polygyny     2.253218   0.971030  2.3204 0.0203170
eboysxp     1.177995   2.577947  0.4570 0.6477065
respatt     4.057854   3.432057  1.1823 0.2370713
landtrans   -5.567216   2.090325 -2.6633 0.0077373
polinteg     3.364141   1.930221  1.7429 0.0813548
socstrat    -5.059963   1.729604 -2.9255 0.0034390

Rho:  -0.0073583 LR test value: 0.22596 p-value: 0.63453
Asymptotic standard error: 0.014268 z-value: -0.51573 p-value: 0.60604
Wald statistic: 0.26598 p-value: 0.60604

Log likelihood:  -358.4752 for lag model
ML residual variance (sigma squared): 184.35, (sigma: 13.578)
Number of observations: 89
Number of parameters estimated: 13
AIC:  742.95, (AIC for lm:  741.18)
LM test for residual autocorrelation
test value:  0.88096 p-value:  0.34794

> pseudoR2<-cor(col.lm$fitted.values,ffd$femsubs)^2
> summary(pseudoR2)
          Min  1st Qu.   Median      Mean  3rd Qu.     Max.
pseudoR2 0.3853  0.3853  0.3853  0.3853  0.3853  0.3853

> df<-data.frame(femsubs=gg$v826,fishimp=gg$v816,huntimp=gg$v817,pathstress=gg$v1260,rainfall =gg$v855,polygyny=gg$v877,eboysxp=gg$v353,fixres=gg$v150,landtrans=gg$v154,polinteg=gg$v157,socstrat=gg$v158)
> df<-data.frame(femsubs=gg$v826,fishimp=gg$v816,huntimp=gg$v817,pathstress=gg$v1260,rainfall =gg$v855,polygyny=gg$v877,eboysxp=gg$v353,denspop=gg$v156,landtrans=gg$v154,polinteg=gg$v157,socstrat=gg$v158)
> kk<-as.matrix(df)
> oo<-matrix(1,length(kk[1,]),1)
> tm<-kk%*%oo
> rr<-which(tm[,1]!="NA")
> wmat<-mat2listw(lds[rr,rr])
> ffd<-df[rr,]
> library(spdep)
> col.lm<-
> lagsarlm(femsubs~fishimp+huntimp+pathstress+rainfall+polygyny+eboysxp+denspop+landtrans +polinteg+socstrat, data=ffd,wm,quiet=FALSE)

Spatial lag model
Jacobian calculated using neighbourhood matrix eigenvalues
Computing eigenvalues ...

(eigen) rho:  -0.6241835  function value:  -1147.631
(eigen) rho:  -0.3808565  function value:  -1006.085
(eigen) rho:  -0.2304722  function value:  -900.5217
(eigen) rho:  -0.1375295  function value:  -816.9473
(eigen) rho:  -0.0800878  function value:  -757.791
(eigen) rho:  -0.04458688  function value:  -724.902
(eigen) rho:  -0.02264609  function value:  -711.2007
(eigen) rho:  -0.009085945  function value:  -707.0115
(eigen) rho:  0.001512760  function value:  -706.5854
(eigen) rho:  -0.001979506  function value:  -706.4188
(eigen) rho:  -0.002161994  function value:  -706.4187
(eigen) rho:  -0.002064636  function value:  -706.4187
(eigen) rho:  -0.002064241  function value:  -706.4187
(eigen) rho:  -0.002064388  function value:  -706.4187
(eigen) rho:  -0.002064383  function value:  -706.4187
(eigen) rho:  -0.002064393  function value:  -706.4187
(eigen) rho:  -0.002064388  function value:  -706.4187
> summary(col.lm)

Call:lagsarlm(formula = femsubs ~ fishimp + huntimp + pathstress + rainfall + polygyny + eboysxp + denspop + landtrans + polinteg + socstrat, data = ffd, listw = wmat, quiet = FALSE)

Residuals:
Min 1Q Median  3Q    Max
-29.585572 -10.223679  -0.059683   8.864702  38.209418

Type: lag

Coefficients: (asymptotic standard errors)

|        | Estimate | Std. Error | z value | Pr(>|z|) |
|--------|----------|------------|---------|----------|
| (Intercept) | 61.573447 | 15.751933 | 3.9089  | 9.270e-05 |
| fishimp | -0.181646 | 0.065300 | -2.7817 | 0.0054073 |
| huntimp | -0.419998 | 0.081808 | -5.1340 | 2.837e-07 |
| pathstress | -1.440055 | 0.420787 | -3.4223 | 0.0006210 |
| rainfall | -2.191237 | 0.663885 | -3.3006 | 0.0009647 |
| polygyny | 1.843448 | 0.676597 | 2.7246  | 0.0064382 |
| eboysxp | 4.154045 | 1.972950 | 2.1055  | 0.0352478 |
| denspop | -1.650607 | 0.950872 | -1.7359 | 0.0825839 |
| landtrans | -2.404004 | 1.133387 | -2.1211 | 0.0339150 |
| polinteg | 2.801116 | 1.379594 | 2.0304  | 0.0423167 |
| socstrat | -2.259084 | 1.106250 | -2.0421 | 0.0411406 |

Rho: -0.0020644 LR test value: 0.10931 p-value: 0.74093
Asymptotic standard error: 0.0058613 z-value: -0.35221 p-value: 0.72468

Wald statistic: 0.12405 p-value: 0.72468

Log likelihood: -706.4187 for lag model
ML residual variance (sigma squared): 187.79, (sigma: 13.704)
Number of observations: 175
Number of parameters estimated: 13
AIC: 1438.8, (AIC for lm: 1436.9)
LM test for residual autocorrelation
test value: 0.2453 p-value: 0.6204

> pseudoR2<-cor(col.lm$fitted.values,ffd$femsubs)^2
> summary(pseudoR2)

Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.3262  0.3262  0.3262  0.3262  0.3262  0.3262
> df<-data.frame(femsubs=gg$v826,fishimp=gg$v816,huntimp=gg$v817,pathstress=gg$v1260,rainfall=gg$v855,polygyny=gg$v877,eboysxp=gg$v353,denspop=gg$v165,landtrans=gg$v154,bodcontctei=gg$v27,socstrat=gg$v158)
> kk<-as.matrix(df)
> oo<-matrix(1,length(kk[,1]),1)
> tm<-kk%*%oo
> rr<-which(tm[,1]!="NA")
> wmat<-mat2listw(lds[,rr])
> ffd<-matrix(1,length(kk[,1]),1)
> library(spdep)
> col.lm<-lagsarlm(femsubs~fishimp+huntimp+pathstress+rainfall+polygyny+eboysxp+denspop+landtrans+bodcontctei+socstrat, data=ffd, wmat, quiet=FALSE)

Spatial lag model
Jacobian calculated using neighbourhood matrix eigenvalues
Computing eigenvalues ...

(eigen) rho:  -0.6214061  function value:  -658.1164
(eigen) rho:  -0.3760662  function value:  -575.1364
(eigen) rho:  -0.2244378  function value:  -453.6262
(eigen) rho:  -0.1307263  function value:  -476.5343
(eigen) rho:  -0.07280941  function value:  -453.6262
(eigen) rho:  -0.0370148  function value:  -443.3539
(eigen) rho:  -0.01489252  function value:  -439.6948
(eigen) rho:  0.01344499  function value:  -439.4594
(eigen) rho:  0.002621023 function value:  -438.7645
(eigen) rho:  0.0002793391 function value:  -438.7676
(eigen) rho:  0.001582269 function value:  -438.7606
(eigen) rho:  0.001618128 function value:  -438.7606
(eigen) rho:  0.001625812 function value:  -438.7606
(eigen) rho:  0.001626133 function value:  -438.7606
(eigen) rho:  0.001626128 function value:  -438.7606
(eigen) rho:  0.001626138 function value:  -438.7606
(eigen) rho:  0.001626133 function value:  -438.7606

> summary(col.lm)

Call: lagsarlm(formula = femsubs ~ fishimp + huntimp + pathstress + rainfall + polygyny + eboysxp + denspop + landtrans + bodcontctei + socstrat, data = ffd, listw = wmat, quiet = FALSE)

Residuals:
     Min      1Q    Median      3Q     Max
-31.56635 -9.06779  -0.33532   7.94368 32.84252

Type: lag
Coefficients: (asymptotic standard errors)

                Estimate Std. Error z value  Pr(>|z|)
(Intercept)     54.394545  18.338676  2.9661  0.003016
fishimp        -0.181969   0.087233 -2.0860  0.036977
huntimp        -0.495474   0.107736 -4.5990 4.246e-06
pathstress     -1.382850   0.526253 -2.6277  0.008596
rainfall       -1.652613   0.799066 -2.0682  0.038623
polygyny       0.945313   0.824890  1.1460  0.251801
eboysxp        7.524213   2.548917  2.9519  0.003158
denspop        -2.103719  1.219341 -1.7253  0.084475
landtrans      -1.372187  1.316436 -1.0424  0.297249
bodcontctei    -0.682382  1.545571 -0.4415  0.658845
socstrat       -2.183064  1.247583 -1.7498  0.080147

Rho: 0.00162616 LR test value: 0.020397 p-value: 0.88643
Asymptotic standard error: 0.0099949 z-value: 0.1627 p-value: 0.87076
Wald statistic: 0.02647 p-value: 0.87076

Log likelihood: -438.7606 for lag model
ML residual variance (sigma squared): 183.59, (sigma: 13.55)
Number of observations: 109
Number of parameters estimated: 13
AIC: 903.52, (AIC for lm: 901.54)
LM test for residual autocorrelation
test value: 0.41257 p-value: 0.52067

> pseudoR2<-cor(col.lm$fitted.values,ffd$femsubs)^2
> summary(pseudoR2)
  Min. 1st Qu.  Median    Mean 3rd Qu.   Max.
  0.3727  0.3727  0.3727  0.3727  0.3727  0.3727
> df<-data.frame(femsubs=gg$v826,fishimp=gg$v816,huntimp=gg$v817,pathstress=gg$v1260,rainfall=gg$v855,polygyny=gg$v877,eboysxp=gg$v353,denspop=gg$v156,respatt=gg$v572,bodcontctei=gg$v27,socstrat=gg$v158)
> kk<-as.matrix(df)
> oo<-matrix(1,length(kk[,1]),1)
> tm<-kk%*%oo
> rr<-which(tm[,1]!="NA")
> wmat<-mat2listw(lds[rr,rr])
> ffd<-df[rr,]
>
> library(spdep)
> col.lm<-lagsarlm(femsubs~fishimp+huntimp+pathstress+rainfall+polygyny+eboysxp+denspop+respatt+bodcontctei+socstrat, data=ffd,wmat,quiet=FALSE)

Spatial lag model
Jacobian calculated using neighbourhood matrix eigenvalues
Computing eigenvalues ...

(eigen) rho:  -0.6201265  function value:  -319.7214
(eigen) rho:  -0.368583  function value:  -278.2213
(eigen) rho:  -0.2131206  function value:  -254.2887
(eigen) rho:  -0.1170395  function value:  -243.7549
(eigen) rho:  -0.05765817  function value:  -240.8551
(eigen) rho:  -0.02491888  function value:  -240.7792
(eigen) rho:  -0.0389931  function value:  -240.6698
(eigen) rho:  -0.03914512  function value:  -240.6698
(eigen) rho:  -0.03901316  function value:  -240.6698
(eigen) rho:  -0.03901225  function value:  -240.6698
(eigen) rho:  -0.03901225  function value:  -240.6698

> summary(col.lm)

Call:lagsarlm(formula = femsubs ~ fishimp + huntimp + pathstress + rainfall + polygyny + eboysxp + denspop + respatt + bodcontctei + socstrat, data = ffd, wmat, quiet = FALSE)

Residuals:
  Min 1Q Median 3Q Max
Type: lag
Coefficients: (asymptotic standard errors)

|                 | Estimate | Std. Error | z value | Pr(>|z|) |
|-----------------|----------|------------|---------|----------|
| (Intercept)     | 78.30261 | 27.73950   | 2.8228  | 0.004761 |
| fishimp         | -0.14894 | 0.10531    | -1.4142 | 0.157299 |
| huntimp         | -0.49520 | 0.14894    | -3.3248 | 0.000885 |
| pathstress      | -0.69085 | 0.73500    | -0.9399 | 0.347250 |
| rainfall        | -0.36871 | 1.07358    | -0.3434 | 0.731267 |
| polygyny        | 1.73732  | 1.24258    | 1.3982  | 0.162067 |
| eboysxp         | 5.99237  | 3.41752    | 1.7534  | 0.079528 |
| denspop         | -2.28637 | 1.51408    | -1.5101 | 0.131024 |
| respatt         | 7.42570  | 4.24252    | 1.7503  | 0.080066 |
| bodcontctei     | -1.30172 | 1.98536    | -0.6557 | 0.512044 |
| socstrat        | -5.56993 | 1.78633    | -3.1181 | 0.001820 |

Rho: -0.039012 LR test value: 1.7591 p-value: 0.18474
Asymptotic standard error: 0.027919 z-value: -1.3973 p-value: 0.16232
Wald statistic: 1.9525 p-value: 0.16232

Log likelihood: -240.6698 for lag model
ML residual variance (sigma squared): 176.17, (sigma: 13.273)
Number of observations: 60
Number of parameters estimated: 13
AIC: 507.34, (AIC for lm: 507.1)
LM test for residual autocorrelation
test value: 1.3552 p-value: 0.24436

> pseudoR2<-cor(col.lm$fitted.values,ffd$femsubs)^2
> summary(pseudoR2)

Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.4588  0.4588  0.4588  0.4588  0.4588  0.4588
References


