

The *Sociological Methodologist*

Newsletter of the Methodology Section, American Sociological Association
Summer 2008

Chair: Ross M. (“Rafe”) Stolzenberg *Chair-Elect:* Tim Futing Liao *Secretary/Treasurer:* Guang Guo
Council: Kenneth C. Land, Pamela Paxton, Aimee Dechter, Trond Petersen, Lingxin Hao, and David B.

Grusky

Newsletter Editor: Lawrence E. Raffalovich

Assistant Editor: David P. Armstrong

Contents

The Legacy of Otis Dudley Duncan and the Contributions Guido Imbens: Introductory Remarks to the Duncan Memorial Lecture, Annual Meetings of the American Sociological Association, August 1, 2008.....	1
Curvature and Interaction in Regression.....	2
Frequently Asked Questions on Structural Equation Models [FAQs on SEMs].....	6
Dataverse Project	8
New Data Initiative in Europe	9
Retirement Announcement	9
Book Announcements.....	10
In the News	10
From the Editor.....	10

The Legacy of Otis Dudley Duncan and the Contributions Guido Imbens: Introductory Remarks to the Duncan Memorial Lecture, Annual Meetings of the American Sociological Association, August 1, 2008

by

Ross M. Stolzenberg
rstolzen@uchicago.edu

This session honors the memory of Otis Dudley Duncan, and offers his survivors a chance to do what he did so often, so well and to such great advantage: learn something useful from a leader in a

cognate discipline. Certainly one of the most celebrated examples of that was Duncan’s self-exposure to path analysis as developed by the geneticist Sewell Wright. We don’t really use full-blown path analysis anymore, and Dudley certainly showed his knowledge of its limitations way back when. But path analysis was a big improvement in its time, and that was enough for Duncan. Dudley favored incremental improvement, just as he favored incremental model building. Messianic social science was not his dish of tea. Dudley hated baloney, hot air and obfuscation of every variety. He used path analysis to cut the baloney, cool the air and clear the sight lines. He worked on academic problems with practical implications. He seemed to me to embody George Orwell’s declaration in “Why I write,” “I write it,” said Orwell, “because there is some lie I want to expose.” So for Duncan and Orwell alike, those lies were falsehoods about social life that were used to justify unfairness, bigotry, inequality and disadvantage. Duncan’s paper, “Inheritance of Poverty or Inheritance of Race” is a fine example. So is, “Is the Intelligence of the General Population Declining?” The answer, by the way, is No, it is not.

It is true that Duncan hated mistakes. But he expected them, and that’s why he spent so much time looking for his own errors. And he forgave other people their mistakes, though he may not have pardoned himself very often. But intentional errors, wanton carelessness and meanness of any sort were not things that Duncan seemed to forgive, perhaps ever. Those failings were taken to be signs of bad character, I think, and portents of trouble to come. Dudley was right. Some of the successful people he found wanting are still around, still successful, still

causing pain and suffering for those near to them, and still writing papers that are wrong the day they are written, and written only because they are useful to those whose motives are narcissistic, ulterior, or both. Such papers make for bad social policy.

Some of Dudley's students went on to great success because of him. Some have been merely incrementally better researchers than they would have been without him. They all contributed to the progress of research and the practical benefits it brings to actual human beings.

Now comes Guido Imbens, and he is, I do believe, the proper scholar to lecture in the name of Dudley Duncan. I am grateful to Guido for agreeing to give this lecture. Guido Imbens is professor of economics at Harvard University. He was previously professor of economics at UC Berkeley, and UCLA, and he did some time at Arizona State and Florence. His graduate study was at Brown University, and his university training was at Erasmus University in the Netherlands.

Imbens lists his interests as "causality, program evaluation, identification, Bayesian methods, semiparametric methods, and instrumental variables." That is a fair summary of topics, but it fails to enlighten about the luxuriant elegance and handy practicality of Imbens' work, even as it summarizes its nominal subjects. Imbens is one of the few who understands the inseparability of modern existentialist despair over identification problems, from modern positivistic euphoria that grows out of Bayesian thinking. The identification problem is that it is tough to pose questions so that they are answerable at all. The Bayesian insight is that the world is fundamentally predictable, and it is going to go on working largely, but not completely, as it has. From this combination, we get new confidence in the power of research to answer questions about what makes the world a better place, and what makes it worse. If I understand Guido correctly, good policy makes the world better, and we need good research and good research methods to make good policy.

Imbens' name has become associated with matching methods, among other nifty things. Guido's work on these methods is not just famous; it's justly famous. It's useful too. He still has a web site at Berkeley, and you can download statistical

programs from it for matching analysis. Matching methods are sometimes presented by others with messianic enthusiasm, promises to cleanse our analyses of the sins of self-selection, and faith in the power of these methods to redeem our analyses through the mystery of identification. What I like about Imbens' approach is that he seems to conceive of it with neither false hopes nor despair. He seems to see his contributions as incremental improvement, something to add to the tools we already choose to use, rather than a decision to lose the entire past. I would say that's the true Bayesian in Guido. I admire that Bayesian very much.

Finally, I want to thank Guido Imbens for his integrity. Some years ago, I edited a journal for the American Sociological Association. Like a number of editors of other academic journals in recent years, I was pressured to suppress publication of forthcoming articles that fairly evaluated the research of several thin-skinned scholars. Many people supported me through that time. I am grateful to them all, and I include Guido Imbens among them. What makes Imbens unique among these good people is that I knew him only from his written work. Yet still, he came forward. So it's a special pleasure to meet him today, and to thank him. Guido Imbens, I look forward to hearing your remarks.

Ross M. Stolzenberg
Professor of Sociology, University of Chicago, and
Chair, Section on Research Methodology, American
Sociological Association

Curvature and Interaction in Regressions

by David F. Greenberg
DG4@nyu.edu

Arthur Stinchcombe (2008) has recently offered advice to researchers estimating a linear regression equation containing curvilinearity in the form of a quadratic term among the predictors, or interaction effects in the form of products of the predictors. His advice is to purge the quadratic or interaction terms of the linear effects by residualizing the curvilinear or product terms on them, standardizing the orthogonalized interaction term, and then using

these standardized, residualized scores as independent variables, along with the unstandardized linear terms, to estimate the regression equation. Stinchcombe's rationale for this advice is that the linear terms "have no substantively meaningful metric," while the coefficients for those terms, and all tests for the significance of the coefficients, will be unaltered by the transformation.

In most research contexts, this is bad advice. It is true that this complicated procedure will yield the same significance test for the quadratic or interaction term as would be obtained from the original equation. But since that test could be conducted without going to the extra work, this is hardly a reason for recommending it. In fact, a good deal would be lost by following Stinchcombe's recommendations. I show this for a simple case of curvilinearity with one independent variable and its square as predictors, and then demonstrate that the same analysis holds when there are additional independent variables, and when interaction terms are present. Nothing in my presentation is original; all of the ideas have appeared in textbook and journal article treatments of regression models, but they warrant repeating because many researchers do not understand them.

Consider the prediction equation obtained by regressing dependent variable y on linear and quadratic terms in x :

$$\hat{y} = a + bx + cx^2 \quad (1)$$

The coefficients b and c represent estimated coefficients even though they are not wearing hats, and we assume that the coefficient c is not equal to zero.

The regression coefficients here cannot have their usual meaning. A regression coefficient is supposed to represent the effect of varying one independent variable while holding the others constant. However, one cannot hold x^2 constant while allowing x to vary; nor can one hold x constant, while allowing x^2 to vary freely. There is really just one independent variable here, not two. As we will see below, the linear term can even be transformed away. Nevertheless, even though the individual coefficients do not have meaning taken

by themselves, they do have meaning as part of the entire prediction equation.

Without estimates of all three unstandardized coefficients, researchers will be unable to compute predicted values of y for given values of x . They will be unable to graph the equation. With variables standardized, researchers will also be unable to compare the coefficients estimated from one sample with those from another, or to test them against hypothesized non-zero population values. Even if the only interest lies in the coefficient for the interaction term, the comparison should be made with unstandardized variables. It is the coefficients for the unstandardized variables that normally represent the strength of causal influences. The standardized coefficients depend on both the unstandardized coefficients and on the standard deviations of the variables. Consequently two standardized coefficients could differ merely because the independent variables in one sample have a larger dispersion than those in the other sample. This makes them ill-suited for cross-sample comparisons (Wright, 1976).

With estimates for all three coefficients one can derive substantively interesting conclusions. To see this, note that eq. 1 is algebraically equivalent to the equation

$$\hat{y} = (a - b^2/4c) + c(x + b/2c)^2 \quad (2)$$

This equation has only an intercept, whose value is $a - b^2/4c$, and a quadratic term in the variable $x^* = x + b/2c$. The linear term has been made to disappear. From eq. 2 we see at a glance that \hat{y} has a minimum (if c is positive) or maximum (if c is negative) when $x = -b/2c$, and that the value of \hat{y} at that point is $a - b^2/4c$. Without the values of all three estimates one wouldn't know this.

To illustrate how informative these coefficients can be, consider educational attainment (*educ*) in the 1998 General Social Survey. When regressed against age (A) and age-squared, we obtain the following OLS estimates for males (sample size 1221):

$$educ = 11.143 + .126A - .00151A^2 \quad (3)$$

All coefficients are statistically significant at the .001 level. Eq. 3 reaches a maximum at $A_{max} = .126/2(.00151) = 41.72$ years old. The mean number of years of education at that age is 13.77. The value at any other age is easily computed from eq. 3. For females, the corresponding prediction equation is (sample size 1595):

$$educ = 10.156 + .173A - .00203A^2 \tag{4}$$

This reaches a maximum at 42.93 years; at that age the mean educational attainment is 13.84 years. We thus derive substantively interesting conclusions: on average, women ultimately finish about as much education as men, but they take a little more than an extra year to do so. Someone following Stinchcombe’s advice would learn only that the standardized coefficient for the quadratic term is -.83 for men, and -1.28 for women, which will mean little to most researchers.

This type of algebraic manipulation can be extended to equations with more than one variable. Thus, if there are two independent variables x and w , with both linear and quadratic terms present, as in this equation:

$$\hat{y} = a + b_1x + b_2x^2 + b_3w + b_4w^2 \tag{5}$$

the rearranged equation is

$$\hat{y} = \left(a - \frac{b_1^2}{4b_2} - \frac{b_3^2}{4b_4}\right) + b_2\left(x + \frac{b_1}{2b_2}\right)^2 + b_4\left(w + \frac{b_3}{2b_4}\right)^2 \tag{6}$$

The approach can also be extended to interaction terms defined as the product of two individual terms:

$$\hat{y} = a + b_1x + b_2w + b_3xw \tag{7}$$

Here we appear to have three independent variables, but the appearance is again deceiving. Of the supposed three independent variables, x , w and xw , it is impossible to hold two of them constant while the third varies. Geometrical interpretation is facilitated by noting that eq. 7 is algebraically equivalent to

$$\hat{y} = \left(a - \frac{b_1b_2}{b_3}\right) + b_3\left(x + \frac{b_2}{b_3}\right)\left(w + \frac{b_1}{b_3}\right) \tag{8}$$

which is the equation for a saddle-shaped surface. Note that we could simplify the equation further by defining new variables, $a^* = a - b_1b_2/b_3$, $x^* = x + b_2/b_3$ and $w^* = w + b_1/b_3$. Then the equation becomes

$$\hat{y} = a^* + x^*w^* \tag{9}$$

We have, by a mere translation of the variables - that is, a shifting of their zero-points, eliminated the direct terms in x and w .

Assuming for purposes of exposition that b_3 is positive, \hat{y} will be larger than the value given by the intercept when the x^* and w^* have the same sign, i.e. when both are positive or when both are negative. Thus, in the first and third quadrants, the predicted value of y increases with x^* and w^* . In the second and fourth quadrants, \hat{y} decreases the farther x and w move from the origin. The value of the quantity in the first set of parentheses to the right of the equal sign in eq. 8 is the value of the dependent variable when $x^* = 0$ or $w^* = 0$, which is to say, when $x = -b_2/b_3$ or $w = -b_1/b_3$.

This type of algebraic manipulation can also be utilized with non-linear regression, as can be seen by considering the estimates obtained by Fagan, Wilkerson and Davies (2007) for a mixed effects Poisson regression model with autoregressive covariance structure. The authors predict that homicide rates in a given police precinct will rise in response to homicide rates in the preceding year in surrounding precincts. In this model the rate parameter is depends on a set of independent variables exponentially, i.e. $\lambda = \exp(Bx)$. Measuring time in years, with $time = 0$ in 1985, they obtain these estimates for total homicides:

Table 1. Inward Contagion of Total Homicides, New York City Census Tracts, 1985-2000

Predictor	B	exp(B)	t
intercept	-1.743	0.175	-70.270
time	0.156	1.168	14.800
homicide contagion	0.050	1.052	26.610
% AA population	0.012	1.102	23.520
time*homicide contagion	-0.008	0.992	7.470

Source: Fagan, Wilkerson and Davies (2007: Table 5, Panel B).

It is the coefficients for *time*, *homicide contagion* and their product that interest us. Consequently we ignore the estimates for % African American population, and consider only the other estimates.

Using the notation of eq. 7, $b_1 = .05$, $b_2 = .156$ and $b_3 = -.008$. For reasons that are not clear, Fagan, Wilkinson and Davies indicate that they consider the critical coefficient to be b_3 , and observe that the odds ratio of .992 shows a positive contagion. On this basis, they conclude that homicides in neighborhoods surrounding a given neighborhood raise its homicide rate. Actually, the negative B coefficient and the odds ratio of .992, which is less than 1, show a reduction. This shows that the effect of contagion on homicide rates was getting smaller over time. This is a finding that the authors did not anticipate in their theoretical discussion.

We learn more by casting the results for the exponential in the format of eq. 8 (we drop the term involving % African American):

$$\begin{aligned} & -1.743 + .050\text{contagion} + .156\text{time} - .008\text{time} * \text{contagion} \\ & = -2.718 - .008(\text{contagion} - .156 / .008)(\text{time} - .05 / .008) \\ & = -2.718 - .008(\text{contagion} - 19.5)(\text{time} - 6.25) \end{aligned}$$

This equation shows that when homicide rates in surrounding precincts increase by an amount $\Delta\text{contagion}$, the exponential in the expression for λ changes by $-.008(\text{time} - 6.25) \Delta\text{contagion}$.

Another way of reaching the same conclusion is to rearrange terms in the right-hand side of the first line in eq. 10 so that it reads $-1.743 + .156\text{time} + (.050 - .008\text{time})\text{contagion}$. If one is interested in possible contagion effects, it is the coefficient of *contagion* that is of particular interest (not the

coefficient for the interaction of *contagion* with *time*). This coefficient receives two contributions, the direct, linear effect (.050) and the interaction effect ($-.008\text{time}$). The coefficient for *contagion* changes sign as *time* goes from less than 6.25 years to more than 6.25 years. Consequently, if the model being estimated is correct, an increase in homicide rates in surrounding census tracts increased the homicide rate in the surrounded census tract during the first six years of the period covered by the study (1985-1991), and reduced it in subsequent years.

The explanation of a time-dependent effect of this sort poses a theoretical challenge. It calls for an explanation of why the direction of an effect reverses at a particular point in time. Not having noticed this reversal, Fagan, Wilkerson and Garth failed to take up the challenge. In the absence of any reason for expecting such a reversal, the researcher might want to consider whether the equation being estimated is properly specified. With a different function of time in the equation, such as the square root, the sign reversal during the period studied would not occur.

In both sets of calculations, the estimated coefficients for the quadratic and linear terms (in the equations with curvilinearity) and for the direct and product terms (in the equation involving interaction) were needed to extract useful information from the raw estimates. The paper-and-pencil computations performed here could not have been done by following Stinchcombe's advice to orthogonalize the curvilinear or interaction terms and standardize them.

ACKNOWLEDGMENTS

I am grateful to Valerie West for identifying typographical errors.

REFERENCES

- Stinchcombe, Arthur L. (2008). "A Practical Appendix on Curvature in Regression," *The Sociological Methodologist* (Winter, 2008): 5-6.
- Fagan, Jeffrey, Deanna L. Wilkinson and Garth Davies (2007). "Social Contagion of

Violence.” In David Flannery, A. Vazsonyi and I. Waldman (eds.), *The Cambridge Handbook of Violent Behavior*. New York: Cambridge University Press.

Wright, Gerald C., Jr. (1976). “Linear Models for Evaluating Conditional Relationships,” *American Journal of Political Science* 20: 349-73.

Frequently Asked Questions on Structural Equation Models [FAQs on SEMs]

by
Kenneth A. Bollen
bollen@unc.edu

Over the many years of teaching and working with Structural Equation Models (SEMs) I have come across a number of questions and misunderstandings about SEMs. I thought that it might help to pose some of these questions and to provide responses. For frequent users of SEMs, most of these points will be known. But for the novice or the curious, I hope that these are helpful.

(1) Aren't SEMs a very specialized type of model?

Actually, it is the opposite. SEMs are a very general model. SEM is a multiple equation model that allows for latent variables, multiple measures of each latent variable, random and nonrandom measurement errors. Many common and less common models are special cases of SEMs. For instance, if you assume a single continuous dependent variable, a single indicator of each covariate, and assume no measurement error in all covariates, then a SEM is equivalent to a regression model. Keep the same assumptions, but assume a dichotomous or ordinal dependent variable with a normally distributed disturbance (error) and the general SEM model is equivalent to a dichotomous or ordinal probit model.

If there are multiple observed dependent (“endogenous”) variables and we continue to assume no measurement error, then a SEM becomes

a simultaneous equation model such as is used in econometrics. If we ignore the structural relations among the latent variables and concentrate just on the effects of the latent variables on their corresponding indicators, then SEMs specialize to cover the factor analysis models that have received most attention in psychometrics.

Latent growth curve models, path analysis, panel data models, principal components, and ARMA time-series models are just a few of the other models that are special cases of a SEM.

(2) When should I consider using SEMs?

If you have multiple dependent variables or if you have covariates that contain measurement error, then SEMs could be useful. Or if you wish to test a measurement model that hypothesizes the relation between a set of indicators and a set of latent variables, SEM provides an excellent tool to examine these ideas. SEMs apply whether your observed variables are continuous or categorical. Given the abstract variables in sociology and the measurement error in the indicators and scales that we use, there are many instances where SEMs would be appropriate.

(3) I ignore measurement error in my analyses and feel that this is a conservative strategy in that if I find an effect, it must be really strong to be detected despite the measurement errors.

This belief that random measurement error always reduces the estimated impact of a variable is widespread, but wrong. The fact is that if you have two or more explanatory variables that have even random measurement error, it is difficult to say what the direction of bias will be if the error is ignored. The estimated coefficients of the variables measured with error could be too small, too big, or just right. Even those explanatory variables that have no or negligible measurement error might be biased by the measurement error in other variables. Only when a researcher takes account of measurement error will she know the direction and magnitude of the biases. Considering how frequently researchers ignore measurement error,

these facts raise questions about the accuracy of the estimates reported in many journal articles.

(4) What is the difference between path analysis and SEMs?

Some researchers say that path analysis refers to multiequation models where there is no measurement error. However, if we look at the work of Sewall Wright, the originator of path analysis, we see that he used path analysis in models that included latent variables and measurement error. So path analysis is more general than it usually is portrayed. The path diagram and the path analysis of direct, indirect, and total effects are important parts of a SEM whether or not the model includes latent variables.

(5) I have heard that interactions cannot be used in SEMs. Is that true?

Interactions among latent or observed variables have been discussed in the SEM literature for nearly a quarter of a century. There has been at least one book on the topic and numerous papers. In fact, there still are new procedures being proposed, so a researcher has several possible ways of introducing interactions.

(6) How important is assuming that the observed variables come from a multivariate normal distribution?

The classic derivations of the maximum likelihood estimator of SEM assumed multivariate normality, but for nearly a quarter century we have known that weaker assumptions are possible while maintaining desirable properties. Of course, the consequences will depend on the estimator but for the maximum likelihood estimator the main consequence is the potential inaccuracy of the significance tests. The consistency of the estimator of the parameters is not impacted. In addition, there are alternative estimators and corrected significance tests that apply even with variables from nonnormal distributions. A far more serious consequence is having a model that has structural misspecifications

such as omitted variables or if we ignore measurement error in an observed variable.

(7) Do researchers who use SEMs think that using these models establishes causality among the variables employed?

Anyone who thinks that SEMs or any other model alone can establish causality is mistaken. SEMs can help assess the correspondence between a model and data, but even an excellent fit does not prove the validity of the model. However, we can detect when the model does not fit the data and this raises questions about its correctness. Sewall Wright's comment on path analysis applies to SEMs as well: "...the method of path coefficients [SEMs] is not intended to accomplish the impossible task of deducing causal relations from the values of correlation coefficients" (Wright, 1934, p. 193). Research design, substantive knowledge, and prior research can all support claims of causality. SEMs can help to determine their plausibility and their implications if true, but it cannot deduce causality just from covariances.

(8) Are extremely large sample sizes required to use SEM?

Not necessarily. The quality of the estimates depends on a number of factors in addition to sample size, so it is not possible to give a single cutoff value for sample size. Researchers have used SEMs on a variety of sample sizes ranging from Ns of less than a 100 to 1000s of cases. The estimator, the validity and size of the model, the number of indicators per latent variable, and the magnitude of the R-squares for the endogenous variables can affect performance. For instance, a model with fewer parameters and higher R-squares can work well in smaller sample sizes than one with lower R-squares and numerous parameters. If for a given model and estimator, the sample size is too low, a researcher often will encounter problems with estimation convergence or parameter estimates that do not make substantive sense (e.g., negative variance or correlation greater than one).

(9) How difficult is it to learn SEM software?

The ease of use of SEM software varies across packages and the features that a researcher wants to use. AMOS, for example, allows a researcher to program models by drawing path diagrams. LISREL has two programming languages; one that uses matrices and one that programs one equation at a time in a simple format. EQS and Mplus have formats not too distinct from what you might find in SAS or Stata. Most SEM software distributors have web sites with example programs. An excellent way to learn the software is to find an example similar to the one of interest to you. Then modify the program to conform to your empirical problem.

(10) How can I learn more about SEMs?

There are a dozen or more books on SEMs. Many journals publish SEM applications and theory and there is a journal, *Structural Equation Modeling*, devoted to SEMs. SEMNET is a listserv of over 2000 subscribers who discuss SEMs. Numerous organizations offer workshops on SEMs, some of these announce their courses on SEMNET. Finally, if you are based at a university, it is not unusual to find one or two SEM courses offered on campus in sociology, psychology, marketing, or possibly other departments.

Kenneth A. Bollen
Odum Institute for Research in Social Science
Carolina Population Center
Department of Sociology
University of North Carolina at Chapel Hill

Dataverse Project

from
Gary King
King@Harvard.Edu

For those who have collected research data and made it available to others, it's nice when people thank you. But it would be nicer to receive formal scholarly citation credit and web visibility for your hard work. The Dataverse Network project is designed to get you that credit and visibility.

The idea is to give you a free "dataverse" (your view of the universe of data) -- which is a virtual archive where you can store, permanently preserve, and distribute your data (or list data from other dataverses) with everyone or only those you approve. Your dataverse is branded as yours, with the look and feel of your web site and on your web site, but since it is served out by an installation of the Dataverse Network at Harvard you needn't install any software or hardware. Some other features include:

- Safe and permanent data storage in preservation format branded as yours.
- No need to translate data when statistical software formats change.
- Can be easily re-branded if you move institutions, but either way will never be lost.
- Formal citation credit for your data, including a globally unique identifier and universal numeric fingerprint.
- Establish an unbreakable link between your data and related published work.
- Easy ways for others to find your data and associated scholarship.
- Share your data with everyone, or those who sign your licensing agreement, or only individuals or groups you approve.
- Allow users to subset, recode, and download your data in any format
- Run many advanced statistical methods via a GUI on-line.

An interesting but under-appreciated fact is that if you are at an institution that receives federal funding, and you share research data or put it on your web site without prior IRB approval, you are violating federal regulations (this includes any research data, even that compiled from information in the public domain, from IRB-approved research protocols, or from any other source.). To avoid this problem, the Dataverse Network has automated the IRB data approval process, and so if you have a dataverse in most cases going to the IRB is unnecessary.

For an example, go to my homepage at <http://gking.harvard.edu> and click on dataverse. To

get your own dataverse, go to the IQSS Dataverse Network, <http://dvn.iq.harvard.edu>. For more information on our open source Dataverse Network project, see <http://TheData.org>.

Gary King
David Florence Professor of Government,
Director, Institute for Quantitative Social Science
Harvard University, 1737 Cambridge St,
Cambridge, MA 02138
<http://GKing.Harvard.Edu>
Direct 617-495-2027
Assistant 495-9271
eFax 812-8581

New Data Initiative in Europe

from
Sonja Drobnic
sonja.drobnic@uni-hamburg.de

EDACwowe is the META-DATA-SHELL for European research and policy making in the areas of work and welfare and is of interest to scholars doing comparative research outside of Europe, too. It gives information on and direct links to comparative and national, quantitative and qualitative data on work and welfare and closely related fields.

EDACwowe is an initiative under the EU Framework 6 Network of Excellence "Reconciling Work and Welfare in Europe" (<http://recwowe.vitamib.com>). It is coordinated and supported by the University of Tilburg (The Netherlands) and by the Danish National Centre for Social Research, SFI (Denmark).

EDACwowe Link: <http://www.edacwowe.eu>

Prof. Sonja Drobnic, PhD
Institute of Sociology
University of Hamburg
Allende-Platz 1
D-20146 Hamburg
Germany
Phone + 49 40 42838-4659
<http://www.wiso.uni-amburg.de/index.php?id=1947>
<http://www.cgg.uni-hamburg.de>

Retirement Announcement

Jack Angle has retired from the Economic Research Service (USDA) to work full time on finding new applications of the Inequality Process. A bibliography on the Inequality Process, with links to downloads where available, is on-line at <http://www.inequalityprocess.org>. Semi-popularized introductory articles are identified. This URL will have to be directly opened by your browser or clicked on as a link until the "crawls" of major search engines get around to finding it. Jack can be reached at angle@inequalityprocess.org.

The Inequality Process (IP) is a simple model of competition for wealth in a population. The IP is in the class of interacting particle systems (ips), the class of a mathematical model at the core of statistical mechanics. A successful interacting particle system does what the micro/macro models of social science claim to do. In representing people in a particle system, each person is stripped of all but the most essential traits – and then referred to as a "particle". While there is no clear cut distinction between a particle system and agent models, it is expected that a particle will be simpler than an agent. As a mathematical model, the IP is similar to the stochastic version of the kinetic theory of gases. In the IP particles have pairwise competitive encounters. Wealth is transferred to winners but since the chance of any particle winning an encounter is 50/50, wealth flows in the long term to particles that lose less when they lose. The IP was abstracted from a speculation by Gerhard Lenski that more productive workers lose less in the competition for wealth. Some econophysicists have noticed that the IP explains a perhaps surprisingly wide variety of empirical quantitative patterns in wealth and income statistics with parsimony. It is a simple, inflexible mathematical form with few parameters to estimate. Explanation in the IP is like that of any theory in statistical mechanics: the model's quantitative patterns closely mimic those seen in empirical data. While model parsimony with wide empirical relevance may get the attention of interdisciplinary physicists, these model traits are less likely to excite the interest of social scientists. They may even be experienced by social scientists as uncomfortable. Those unlikely to experience

discomfort with the IP are encouraged to click on <http://www.inequalityprocess.org>.

Book Announcements

from
Patricia Leavy
pleavy@stonehill.edu

Method Meets Art: Arts-Based Research Practice by Patricia Leavy (Guilford, 2009) presents a comprehensive introduction to arts-based research (ABR) practices, which scholars in multiple disciplines are fruitfully using to reveal information and represent experiences that traditional methods cannot capture. Each of the six major ABR genres—narrative inquiry, poetry, music, performance, dance, and visual art—is covered in chapters that introduce key concepts and tools and present an exemplary research article by a leading ABR practitioner. The book discusses the kinds of research questions these innovative approaches can address and offers practical guidance for applying them in all phases of a research project, from design and data collection to analysis, interpretation, representation, and evaluation. The book can be used by scholars interested in ABR or as a teaching tool in courses on interpretive inquiry, narrative inquiry, qualitative methods, and research methods.

The Handbook of Emergent Methods edited by Sharlene Hesse-Biber and Patricia Leavy (Guilford, 2008) comprehensively examines emergent qualitative and quantitative theories and methods across the social and behavioral sciences. Providing scholars and students with a way to retool their research choices, the volume presents cutting-edge approaches to data collection, analysis, and representation. More than 30 leading researchers describe alternative uses of traditional quantitative and qualitative tools; innovative hybrid or mixed methods; and new techniques facilitated by technological advances. Chapters explore the strengths and limitations of each method for studying different types of research questions and offer practical, in-depth examples.

In the News

In mid-July, Reuters published an article reporting on Guang Guo's research linking specific genetic variants to delinquent behavior in young men. The article is available online at <http://www.reuters.com/article/healthNews/idUSN1444872420080714>.

From the Editor

Gary King's description of the Dataverse Project came to us through SOcNET (SOcNET@LISTS.UFL.EDU), the discussion list for INSNA (<http://www.insna.org>). A proposal to establish such a repository, from Jeremy Freese, was published in the Spring, 2006 issue of this newsletter.

We thank Francois Nielsen for sending us the Reuters article describing Guang Guo's research. Unfortunately, I could not obtain permission to reproduce the article. A while back, I followed a reference (in Burnham and Anderson's *Model Selection and Multimodel Inference*) to an article titled "Least Square Fitting of an Elephant" by James Wei in the February, 1975, issue of *Chemtech*. I thought it interesting and thought you might, also. I obtained permission from Professor Wei to reproduce the article. However, Chemtech, the copyright holder, refused. A brief description can be found at <http://www.unc.edu/courses/2006spring/ecol/145/001/docs/lectures/lecture5.htm>.

This is my last issue as editor of *the Sociological Methodologist*. I think it an important service to the Section, but I've been at it too long. I encourage you to volunteer for this responsibility. It is not at all tedious, and requires minimal time commitment. The only hard part is getting folks to submit material. That does take a bit of creativity and some patience, but our members do have much to contribute. I will, of course, do what I can to help out and smooth the transition. If you have questions or concerns about this, let me know (L.raffalovich@albany.edu). Contact Rafe (rstolzen@uchicago.edu) to volunteer.