INTRODUCTION TO STRUCTURAL EQUATION MODELS

I. Description of the course.

A. Objectives and scope of the course.
B. Logistics of enrollment, auditing, requirements, distribution of notes, access to programs.
C. Syllabus: assignments, readings, grading, topics.

II. Historically, the major utility of structural equation models.
A. Multiple equations.
B. Measurement error (multiple indicators).
C. Non-recursive relationships.

II. Brief History of Structural Equation Models (a way of representing phenomena using mathematical linear equations of random variables).

A. Origins: Genetics, Economics, Psychology, Sociology
1. Genetics: Sewall Wright introduced path analysis to population genetics in 20s and 30s.
a) Invented path diagrams: causal arrows, double-headed arrows for correlation, random error terms.
b) Placed path coefficients (standardized regression coefficients) over the arrows.
2. Economics: Econometricians developed systems of simultaneous equations in the forties and fifties to deal with nonrecursive relationships.
a) The Dutch Economist Tinbergen (1936) specified simultaneous equations (used OLS); Koopmans (1950) and Hood and Koopmans (1953) (as part of the Cowles Commission for Research in Economics at the University of Chicago) dealt with estimation (stimulated by Haavelmo 1943).
b) Klein (1950) and Klein and Goldberger (1955) estimated a macro model of the U.S. economy.
3. Sociology: In the early 1960s Blalock, Hodge, Duncan, and Costner started using path analysis in models of status attainment.
a) Landmark paper was Duncan's (1966) "Path Analysis: Sociological Examples," and Blau and Duncan's (1964) American Occupational Structure.
b) Duncan read Sewall Wright's theorems and applied them to sociological examples.
4. Psychology: Psychologists, like Spearman and Thurstone, had developed exploratory factor analysis in the 20s and 30s as a way of identifying dimensions of intelligence.
a) Lawley (1940; see Lawley and Maxwell 1971) developed maximum likelihood estimation of the factor model.
b) Jöreskog (1967; 1969) applied maximum likelihood estimation to "confirmatory" factor analysis and with his programmer Sörbom developed a program to estimate it.
5. At approximately the same time, Jöreskog (1970), Keesling (1972), and Wiley (1973) each developed independently a linear structural equation model that combined confirmatory factor analysis with path analysis. Each began with a covariance matrix of observed variables and then specified a system of structural equations underlying that matrix. Hence the term, "covariance structure model" or "analysis of covariance structures."
6. The system of equations that combined observed variables and "unobservable" or latent variables drawn from factor analysis. They did this by specifying a "measurement" model relating observables to unobservables; and a "structural" model relating unobservables to each other.
7. Jöreskog called his version, the "LISREL" model, with the acronym standing for LInear Structural RELations.

B. In 1970, Karl Jöreskog and his programmer, Dag Sörbom, developed a computer program using maximum likelihood to estimate Jöreskog's covariance structure model. They called it the LISREL program.
   1. Uses maximum likelihood estimation for optimal estimates -- this was a watershed in the field -- integrated path analysis used by sociologists, factor analysis used by sociologists and simultaneous equations used by econometricians in a unified framework.
   2. Can deal with measurement models with latent variables and simultaneous equation models.
   3. The program has been greatly improved over the last three decades; it is now in its eighth version.

C. Major advances in covariance structure models.
   1. Browne's (1974) asymptotic distribution-free GLS estimator, based on the generalized methods of moments.
   2. Muthén's (1984) general method for analyzing measurement models with ordinal, categorical, and censored measures.
   3. Muthén's method for estimating multi-level models (including growth curve models and some hierarchical linear models) using covariance structure models.

D. A number of new computer programs have appeared—each has some different twist or new feature.
   1. Peter Bentler's EQS program uses scalar algebra, first to introduce methods for non-normal distributions.
   2. Bengt Muthén's LISCOMP program was the first to introduce methods for analyzing ordinal and categorical variables. His more recent program, Mplus, provides models for ordinal, categorical, growth curve, multi-level data all within a covariance structure framework. It will also estimate mixture models for latent class and growth models.
   3. Stata 12 has Structural equation modeling (SEM) using either graphical commands (like SIMPLIS) or command syntax in scalar algebra (like EQS), as well as GSEM (Generalized Structural Equation Models) and GLAMM (Generalized Linear Latent and Mixed Models).
   3. R has John Fox's sem package and Yves Rosseel's lavann package.
   4. A few others: Ronald Schoenberg's MILS program in GAUSS; R.P. MacDonald has a program; Glymore et al.'s TETRAD program for exploring causal structures; Arbuckle has AMOS, available through SPSS, SAS has Proc CALIS. There are others as well.

III. Assumptions and Limitations of Linear Structural Equation Models.

A. Linear relationships (and usual assumptions of the general linear model).

B. Large samples: for the general case, the estimator -- maximum likelihood -- relies on asymptotic distribution theory (what happens when the sample approaches infinity) for desireable properties (unbiased and efficient).

C. Interval-level measures.
   1. Most recent versions of programs can handle categorical and discrete measures - initially LISCOMP, now a version is available in LISREL 8 and Mplus.
   2. These methods are restricted to moderate-sized models and very large samples. Models cannot deal with discrete latent variables (Mplus can). We'll try to get to them toward the end of the semester.

D. Multivariate normality.
   1. In general, the estimator and test statistics assume multivariate normal distributions. Some specific models require that only some observed variables are multinormal (e.g., a simultaneous equation and linear regression models only need assume disturbances are distributed multinormal).
   2. LISREL 8 uses weighted least squares (Browne's asymptotically distribution-free generalized least squares estimator), which originally appeared in EQS, and also is used in Mplus, Amos, etc.

3. But even this estimator assumes normally-distributed latent variables underlying non-normally-distributed observed variables. (This is probably reasonable in most applications.)
4. Mplus and GLAMM provides a latent class model, which allows for discrete and ordered latent variables that are not assumed to be normally-distributed.

E. Does not deal with problems of sample selection, nor does it generally deal with violations of assumptions of the general linear model, like serial correlation and heteroskedasticity (except in special situations).

F. Originally ignored sampling weights from large scale complex survey designs, but now Mplus, LISREL 8.8, and Stata SEM can handle sampling weights.

IV. General Classes of Models Estimable Using Covariance Structure Analysis.

A. Single equation linear models (e.g., regression, ANOVA, ANCOVA).
B. Recursive systems of linear equations and path analysis.
C. Seemingly-unrelated regressions.
D. Simultaneous equation (non-recursive) models.
E. Confirmatory factor models (including second-order factor models).
F. Cannonical correlation models.
G. Linear structural equation models with unobserved variables and multiple indicators.
H. Multiple-group models (for modeling interaction effects).
I. Multiple-indicator multiple-cause (MIMIC) models.
J. Models for sibling data and other forms of nested data (certain random effects and fixed effects models); models for datasets with some values missing at random.
K. Models for panel data, including random effects, fixed effects, and autoregressive models.
L. Latent growth curve models and other multi-level structures.

V. Recent Developments.

A. Multi-level models in various forms.
1. Models for siblings and other nested data structures with pairs of contexts.
2. Latent growth curve models.

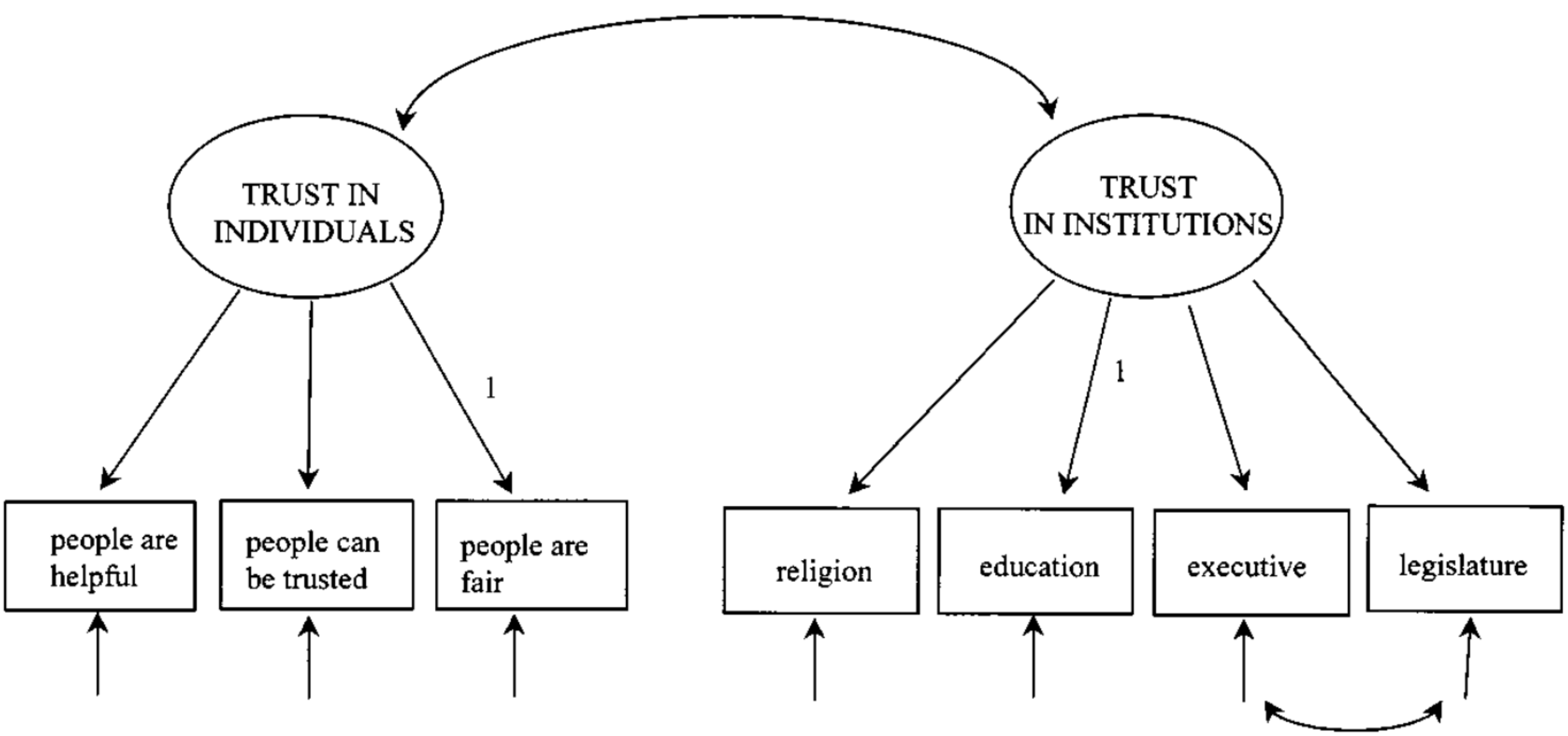
B. The use of Markov chain Monte Carlo methods to estimate complex covariance structure models by Arminger, Muthén, and others.

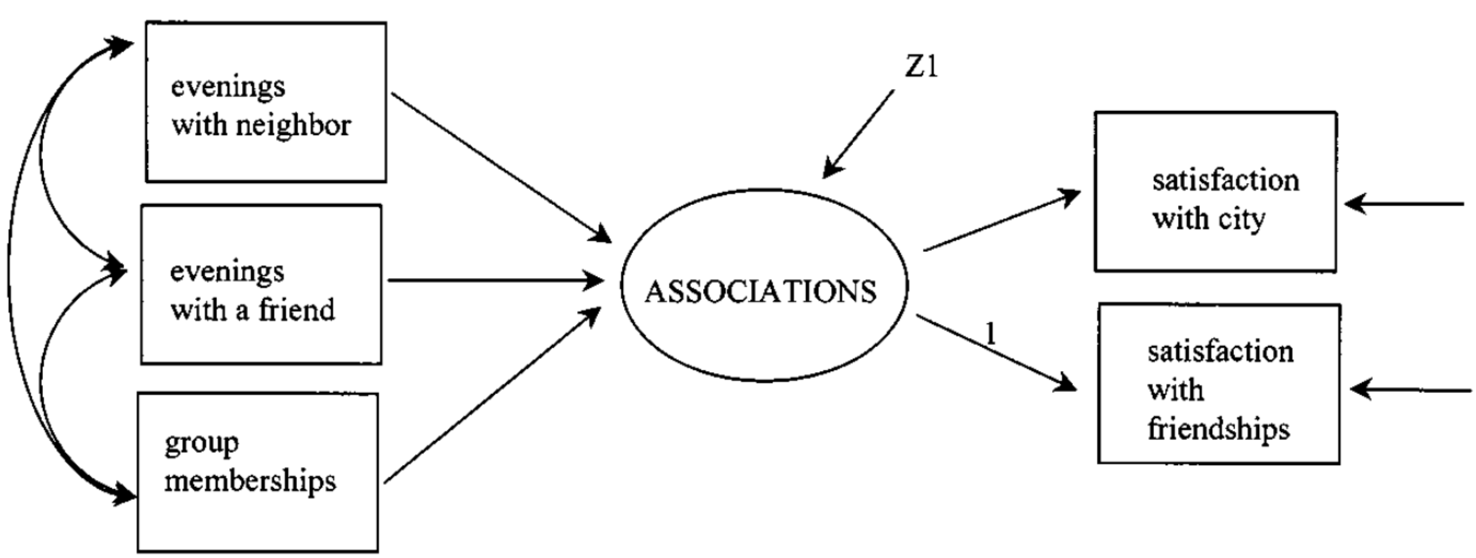
C. Mixture models for latent classes and latent classes of trajectories (Muthén; see also Arminger). These models allow for unobserved heterogeneity in the sample, in which different individuals are members of different subpopulations, but the subpopulation membership is unknown and must be estimated from the data.
1. Latent class model: assign individuals to different latent classes based on binary outcomes.
2. Latent growth mixture model: assign individuals to different latent classes based on estimated trajectories or growth curves.
3. These models depart from the covariance structure framework as they are not estimated from a covariance matrix.

D. Generalized linear models.
1. Jöreskog and Sörbom are incorporating into their LISREL program a generalized linear model—a generalization of the linear model, which includes a systematic component (linear relationship), a random component, and a link function linking the systematic to the random components. The link function (e.g., identity, log, logit, and reciprocal) has associated conditional probability distributions (normal, poisson, binomial, and gamma).
2. Skrondal and Rabe-Hesketh (2004) have developed a framework, generalized latent variable modeling, which integrates generalized linear models, latent variable models, multi-level models, and longitudinal models. Their Stata program is GLLAMM (generalized latent and mixed models). Everything above becomes a special case! See also Stata's GSEM.

E. Recent application of graphical models to causality and structural equations.
   a. Extension of TETRAD II (Spirites et al. 1993) approach, in which underlying structure is discovered through an algorithm that eliminates relationships by identifying causal independencies. Fragments of causality are put together to create causal models.
   b. A Bayesian approach is being applied to the graph-based learning methods.
   c. This material is covered in Judea Pearl's book *Causality*.
   d. Thomas Richardson covers this in his CS&SS course 566, Causal Modeling.

EXAMPLES OF SUBSTANTIVE APPLICATIONS OF SEM





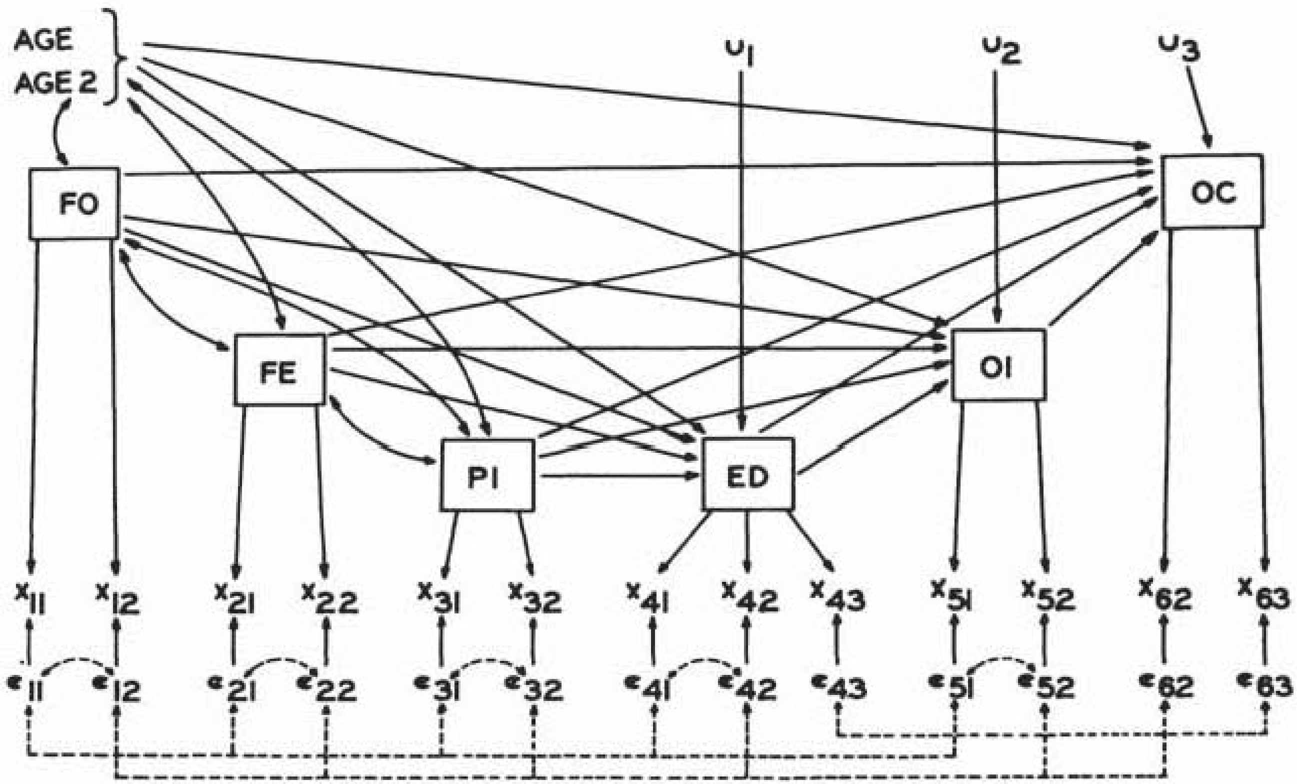


FIG. 2.—A structural equation model of the stratification process with measurement errors. Variables are defined in table 1.

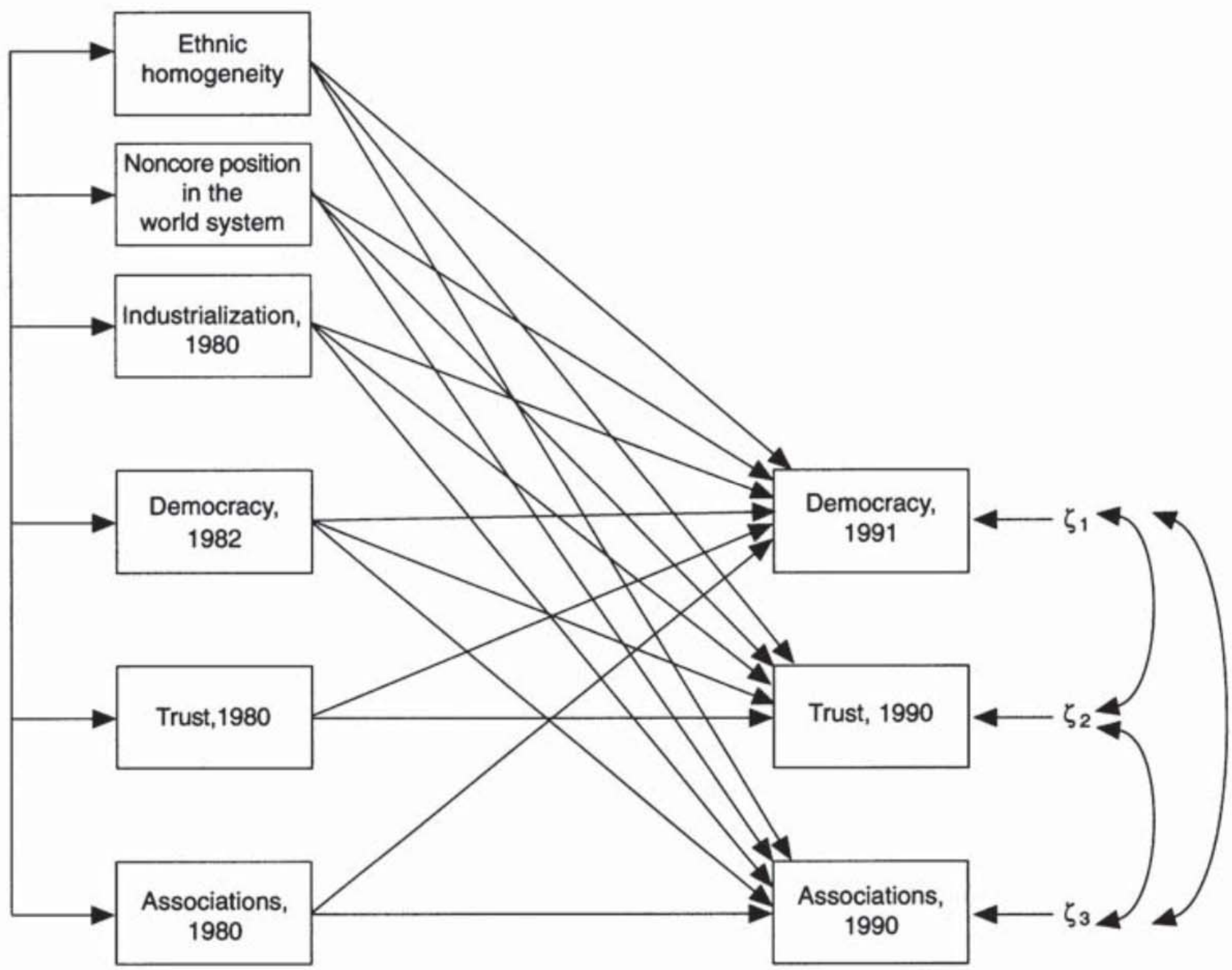


Figure 1. Cross-Lagged Panel Model of Social Capital and Democracy: World Values Survey

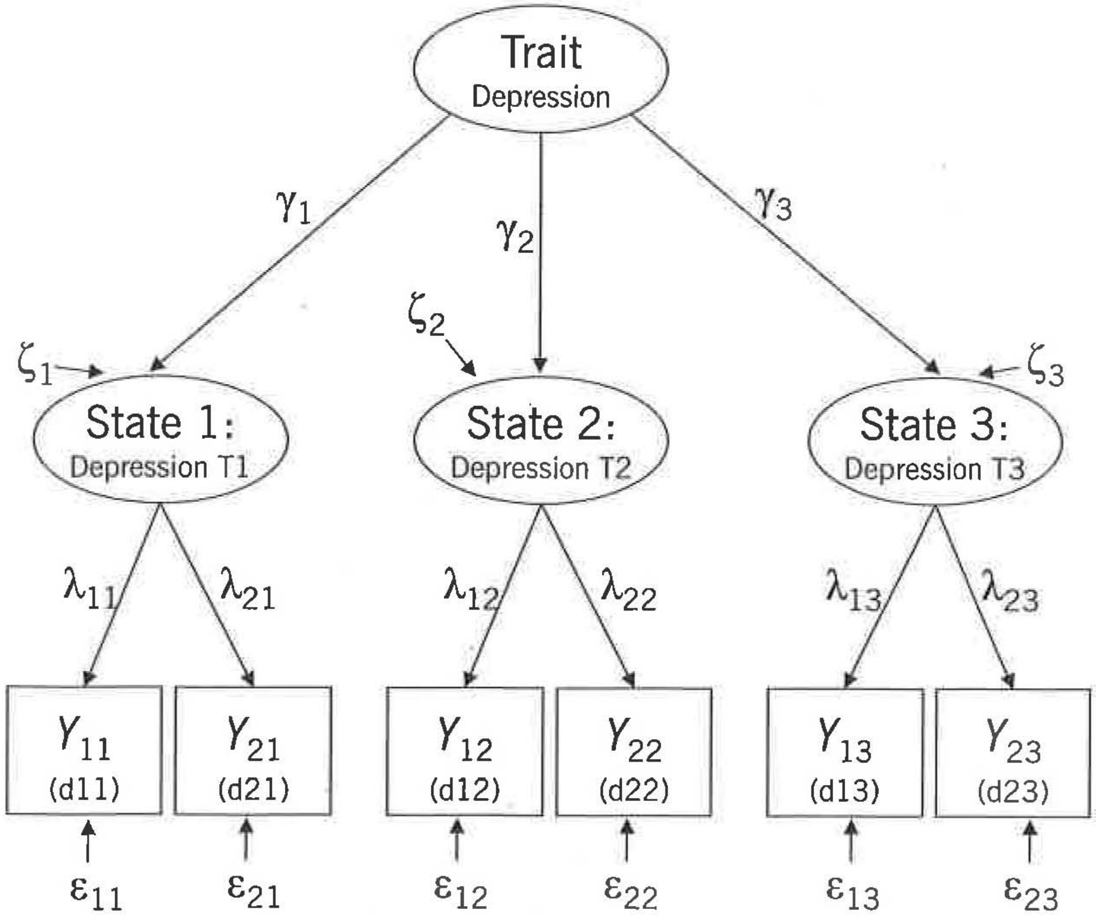


FIGURE 4.11. Latent state–trait (LST) model for depression measured on three occasions of measurement (T1–T3). Y_{ik} = observed variable i measured at time point k; ε_{ik} = measurement error variable; λ_{ik} = first-order factor loading; γ_k = second-order factor loading; ζ_k = LS residual variable. The variable names of the indicators used in Mplus appear in parentheses.

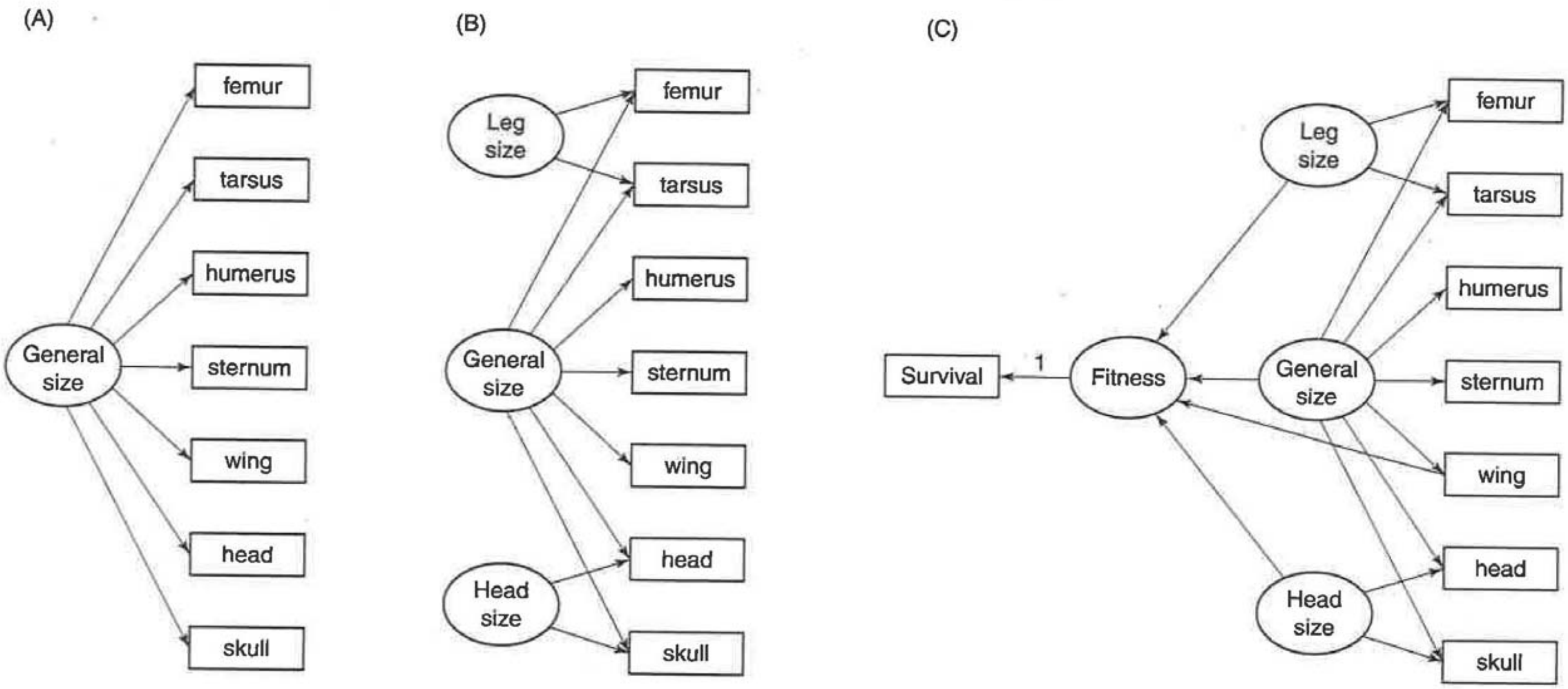
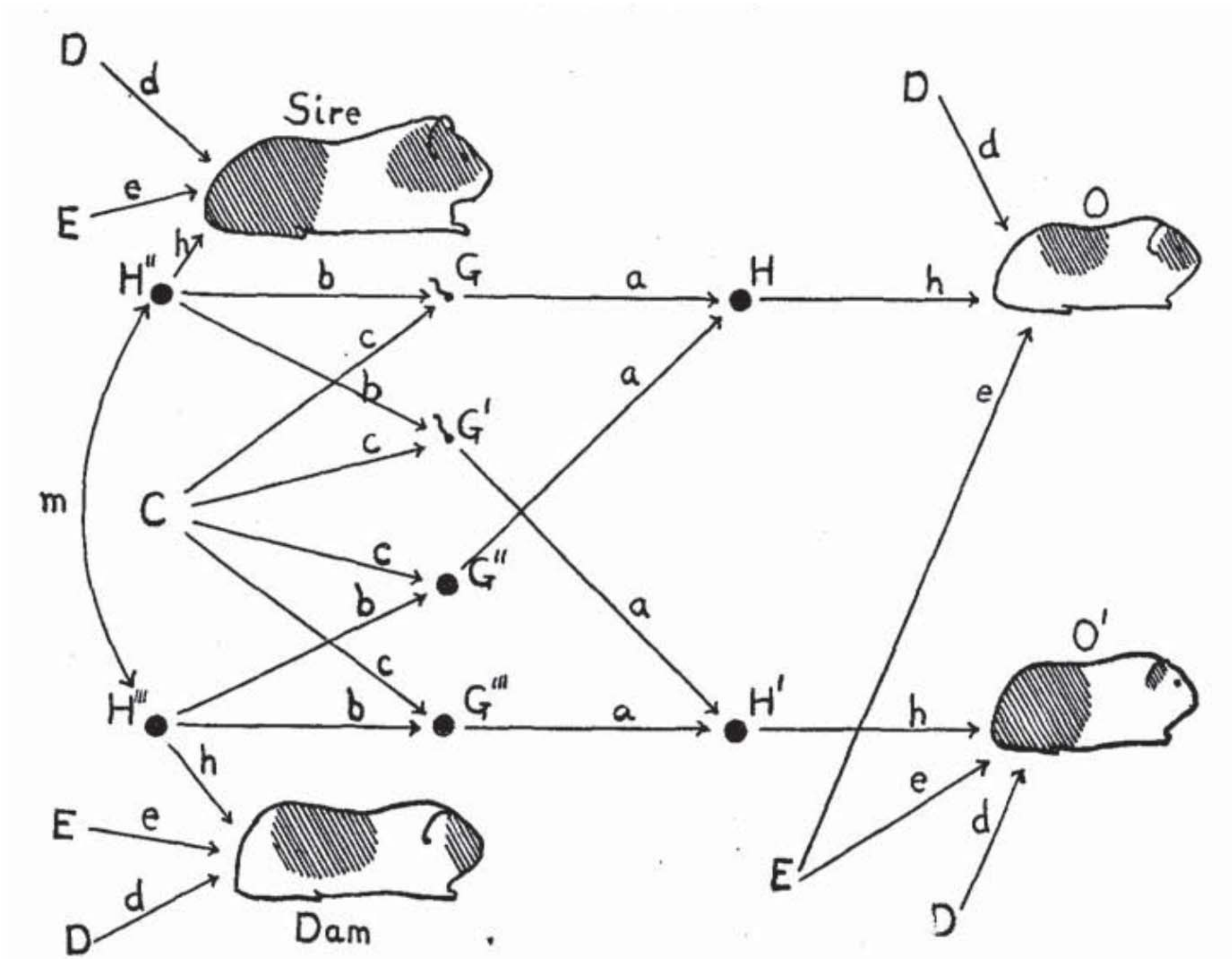


Figure 6.11. Pugesek and Tomer's model of the classic Bumpus House Sparrow data.


Sire
Dam
O
O'
D
d
E
e
h
H''
H'''
H
H'
G
G'
G''
G'''
C
m
a
b
c

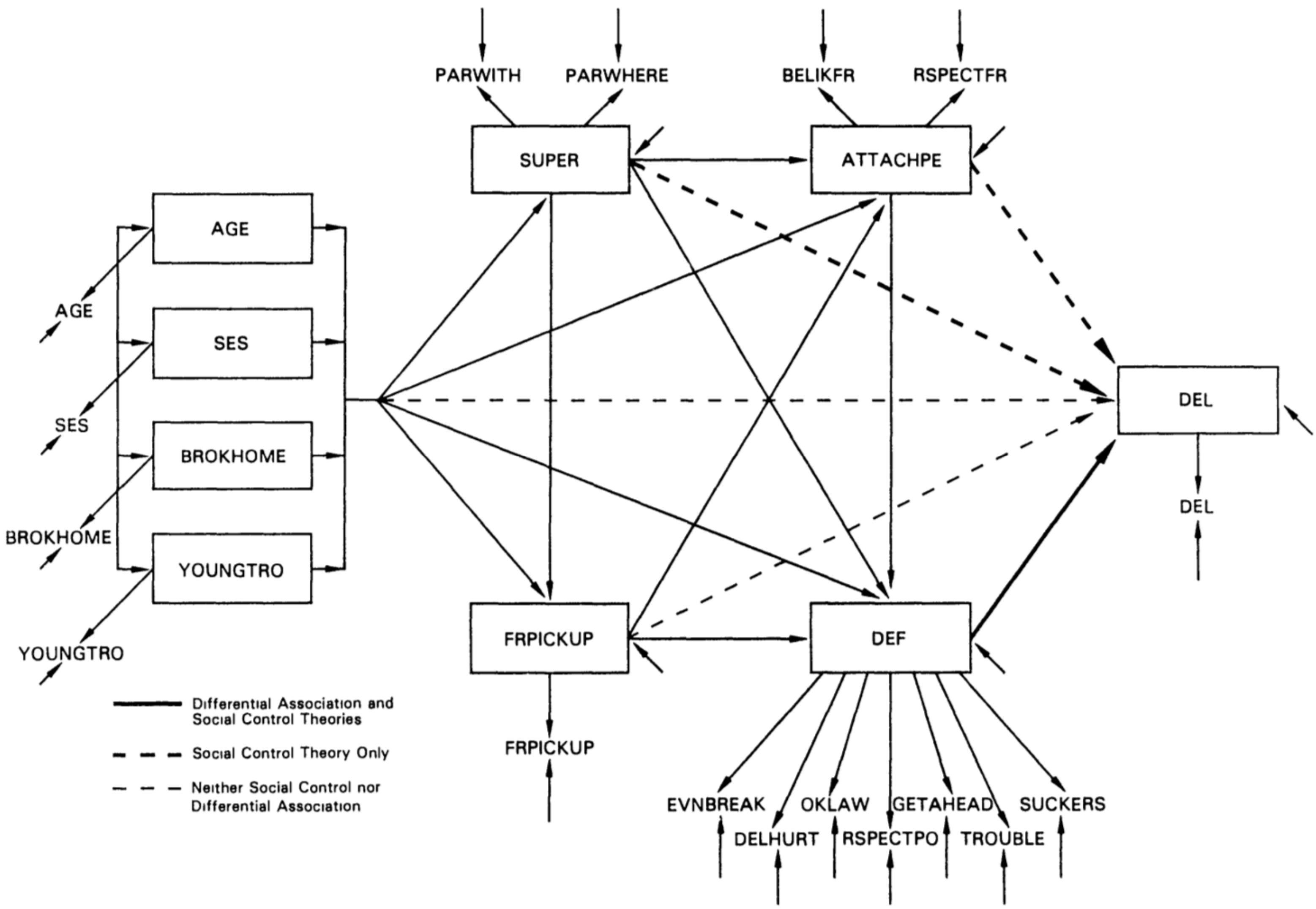


Fig. 1. Path Diagram of the Full Structural Equation Model of Delinquency

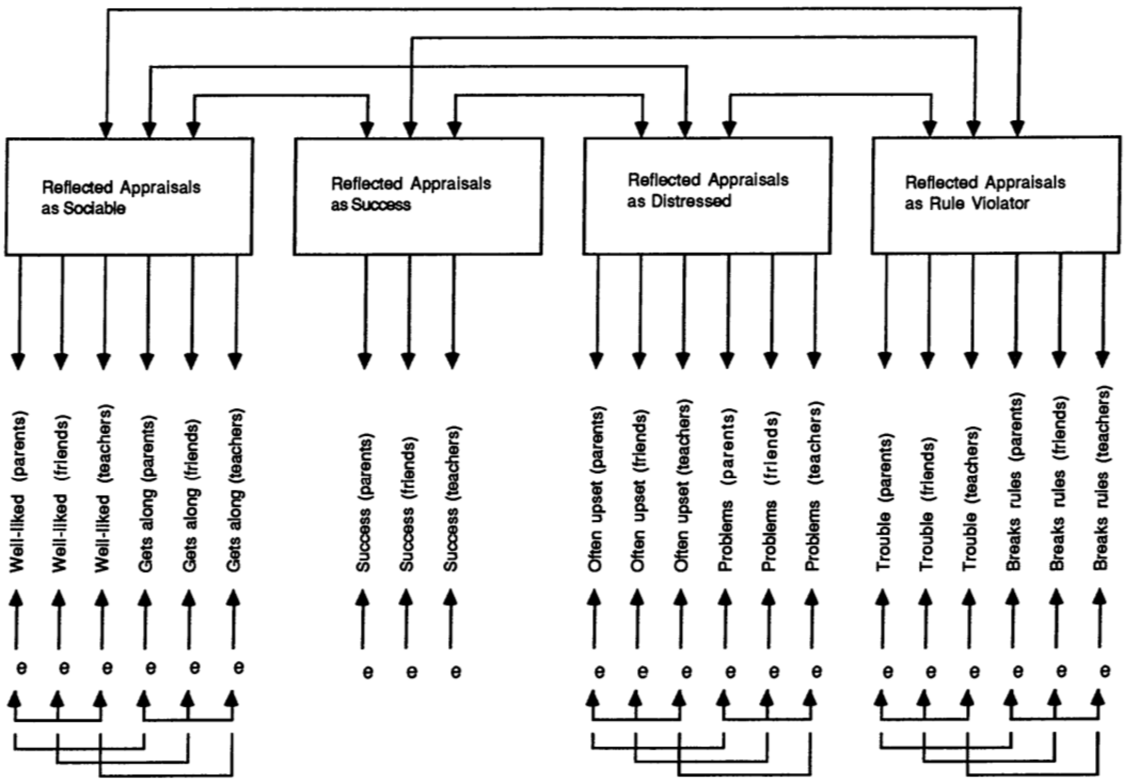


FIG. 2.—A measurement model of reflected appraisals of self

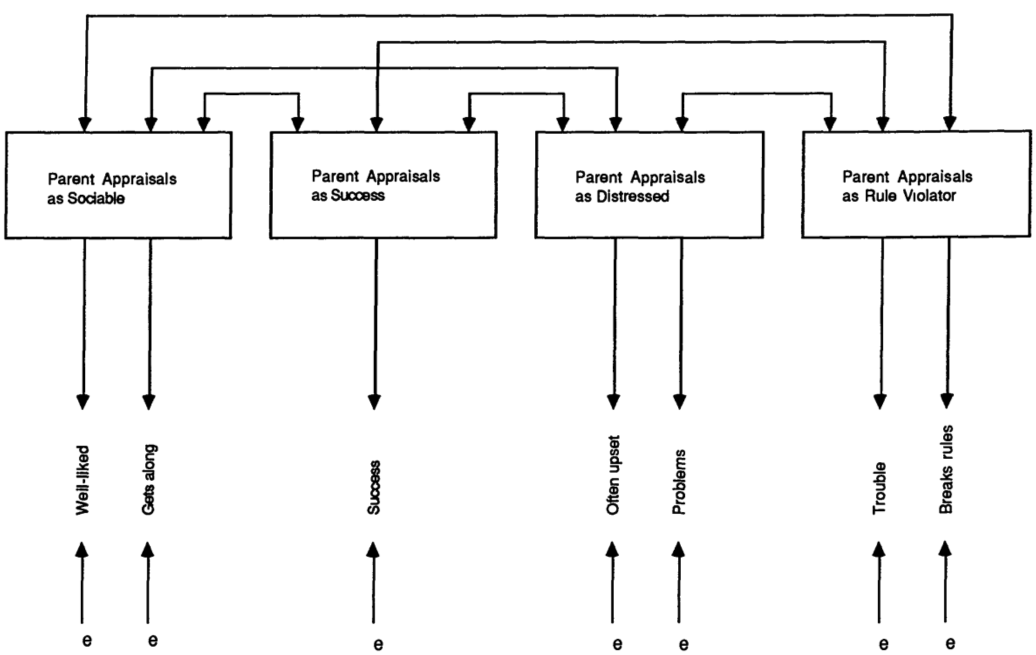


FIG. 3.—A measurement model of parental appraisals of youths

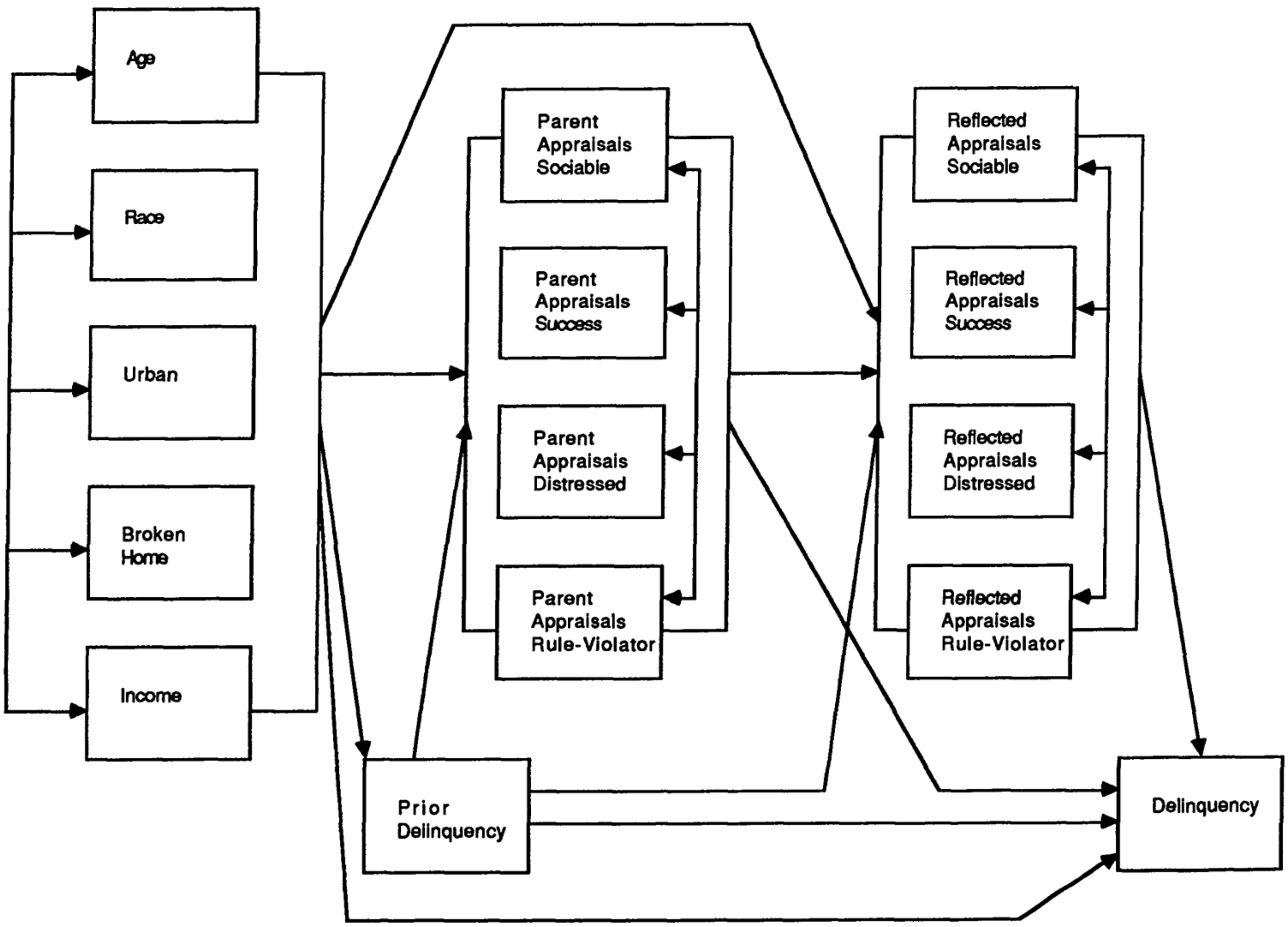
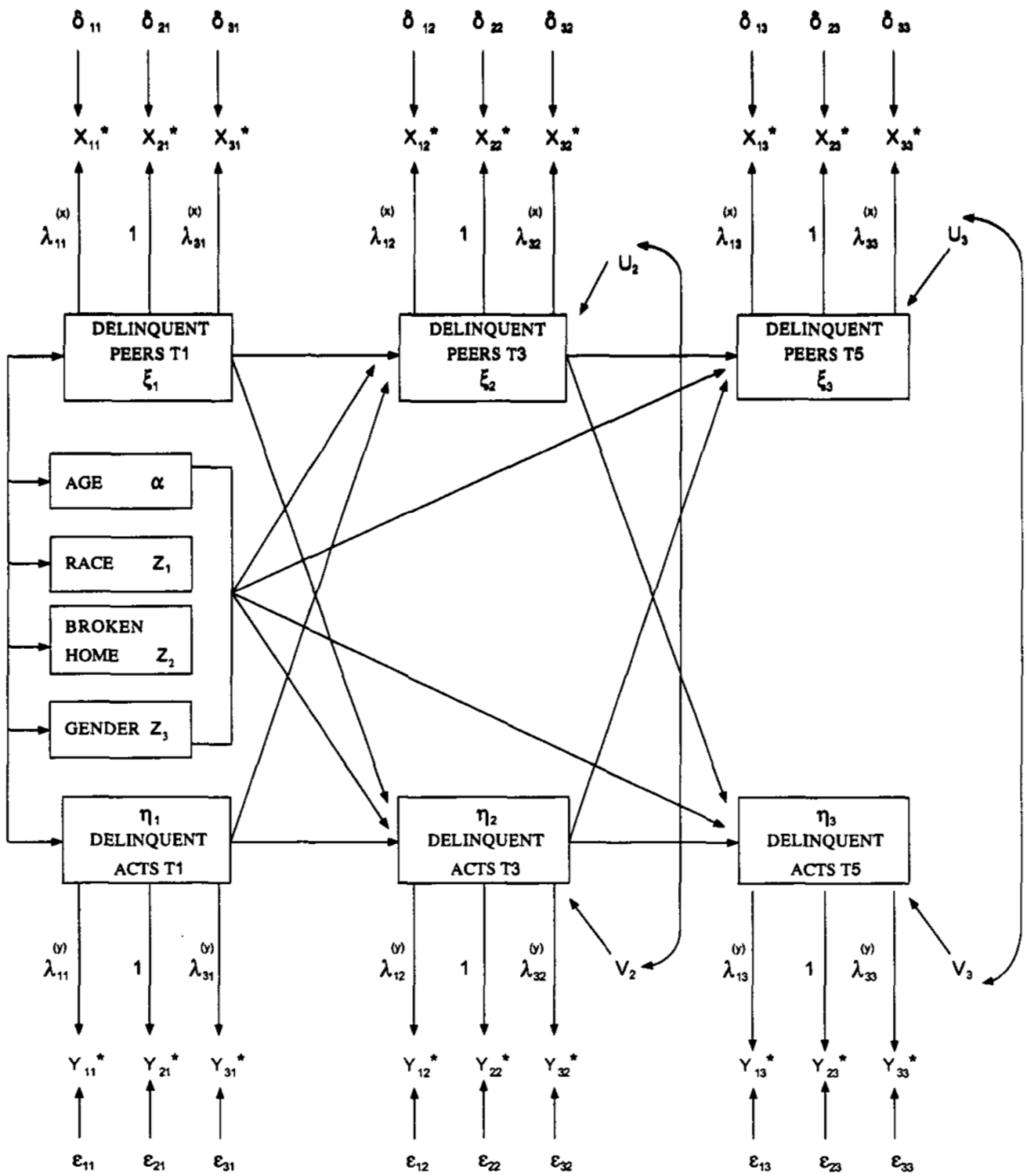


FIG. 4.—A substantive model of parental appraisals, reflected appraisals, and delinquency.

Figure 4. Path Model of a Reciprocal Causal Relationship of Property Delinquent Peers and Property Delinquency



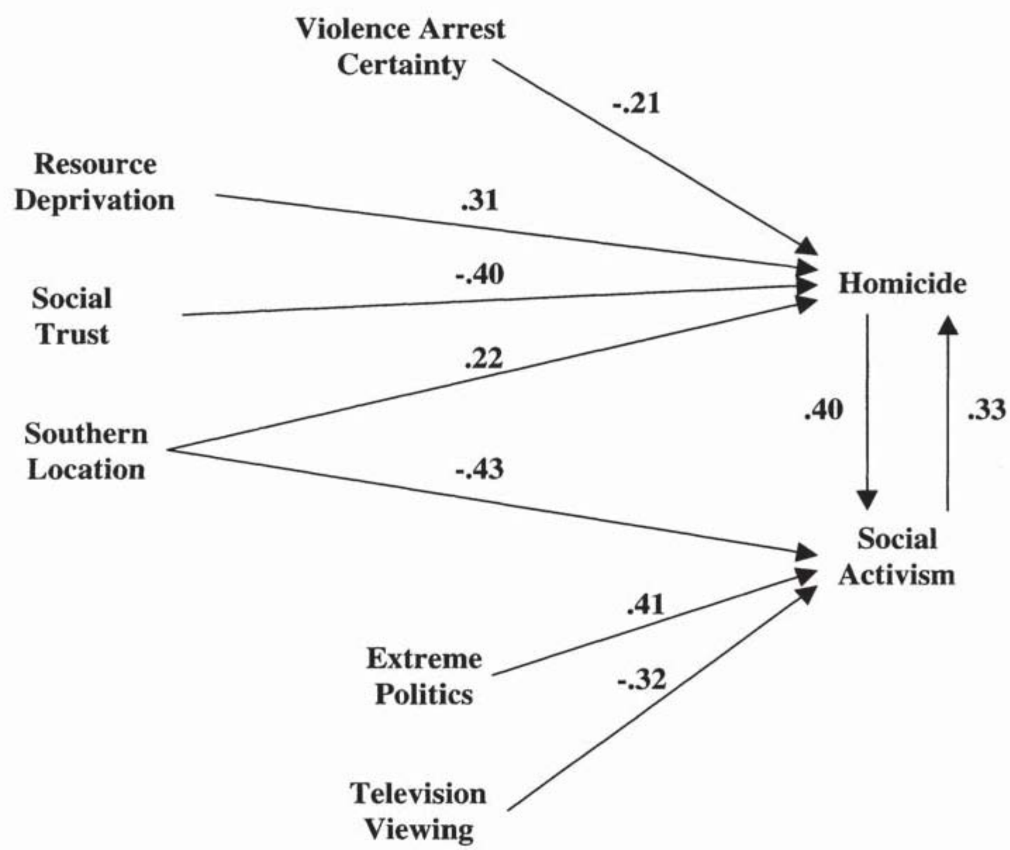


Figure 2. Relationship between Social Activism and Homicide Rates, Reduced Form Model
Note: Model Fit Statistics: $\chi^2 = 13.99$, 17 df ($p = .67$), RMSEA = .000, CFI = 1.00, GFI = .933; Adjusted R^2: Homicide (.859), Social Activism (.557). Standardized estimates are displayed. For clarity of presentation, correlations between exogenous variables are omitted. RMSEA = root mean square error of approximation; CFI = comparative fit index; GFI = goodness-of-fit index.

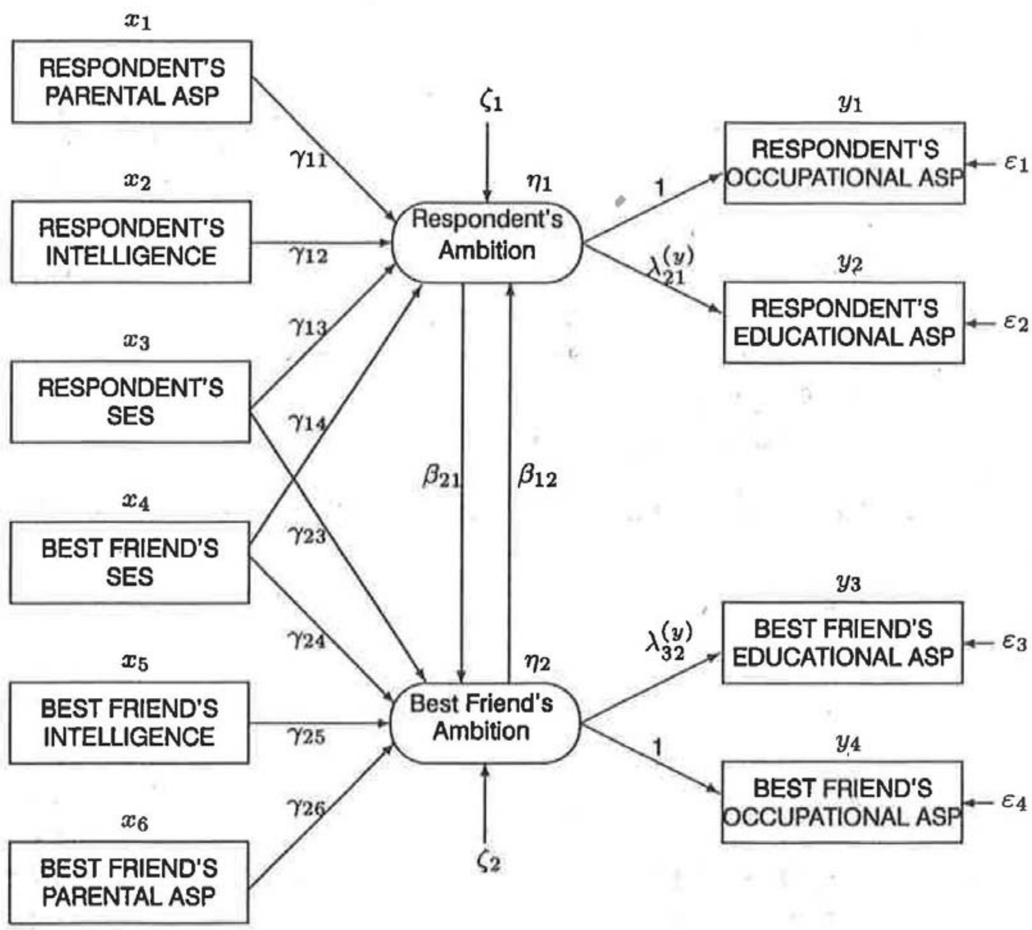


Figure 5.6 Path Diagram for Peer Influences on Ambition

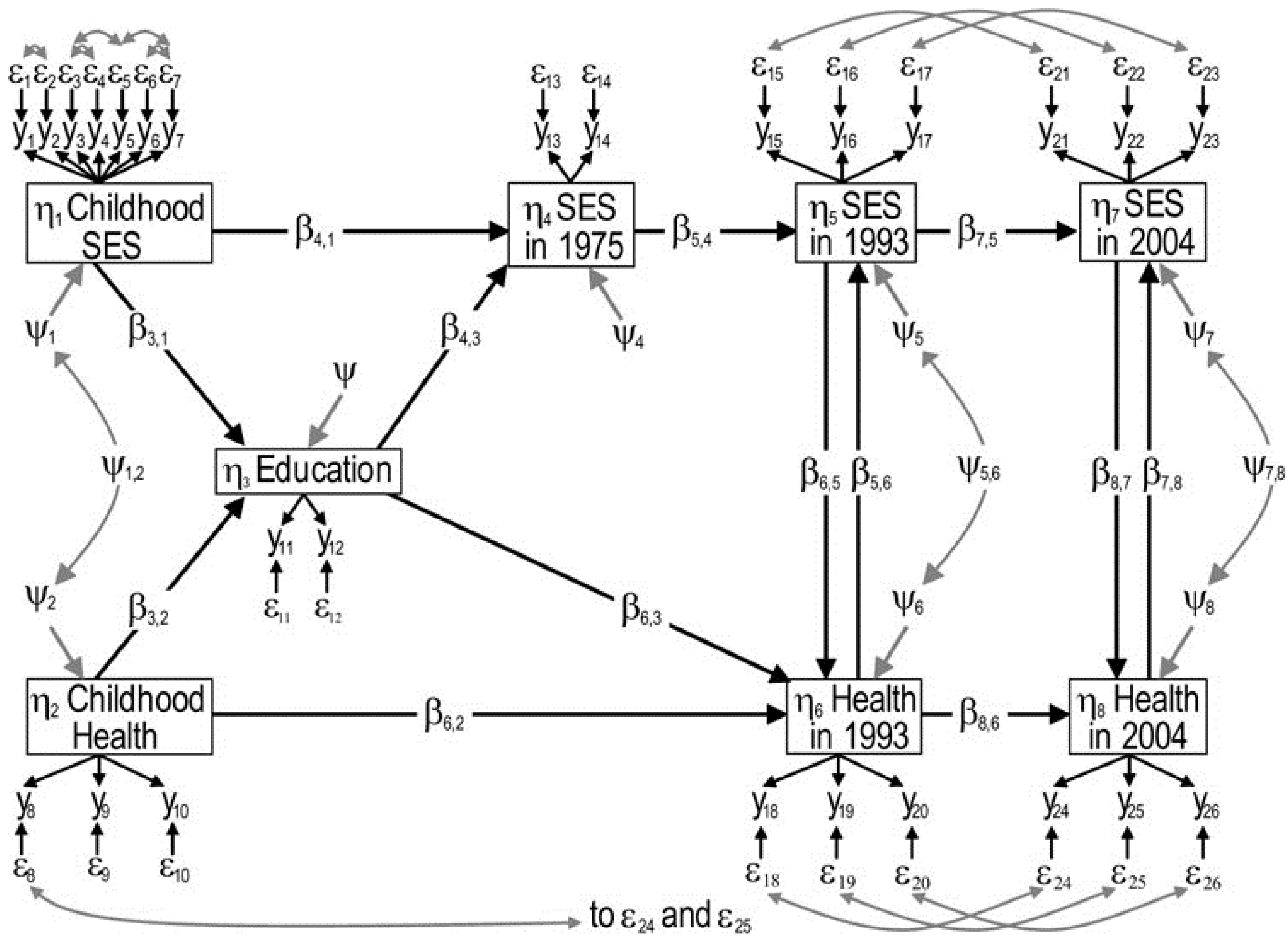

ε1 ε2 ε3 ε4 ε5 ε6 ε7
y1 y2 y3 y4 y5 y6 y7
η1 Childhood SES
ε13 ε14
y13 y14
η4 SES in 1975
ε15 ε16 ε17 ε21 ε22 ε23
y15 y16 y17 y21 y22 y23
η5 SES in 1993
η7 SES in 2004
β4,1
β5,4
β7,5
ψ1
β3,1
β4,3
ψ4
ψ5
ψ7
ψ
ψ1,2
η3 Education
β6,5 β5,6
ψ5,6
β8,7 β7,8
ψ7,8
y11 y12
ε11 ε12
ψ2
β3,2
β6,3
ψ6
ψ8
η2 Childhood Health
β6,2
η6 Health in 1993
β8,6
η8 Health in 2004
y8 y9 y10
ε8 ε9 ε10
y18 y19 y20 y24 y25 y26
ε18 ε19 ε20 ε24 ε25 ε26
to ε24 and ε25

Figure 4. A path model for the resemblance of impulsivity and antisocial

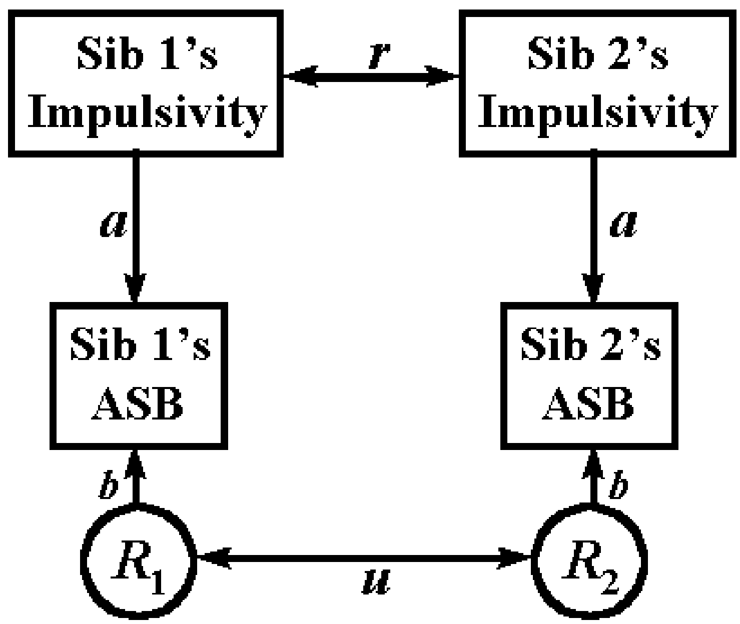
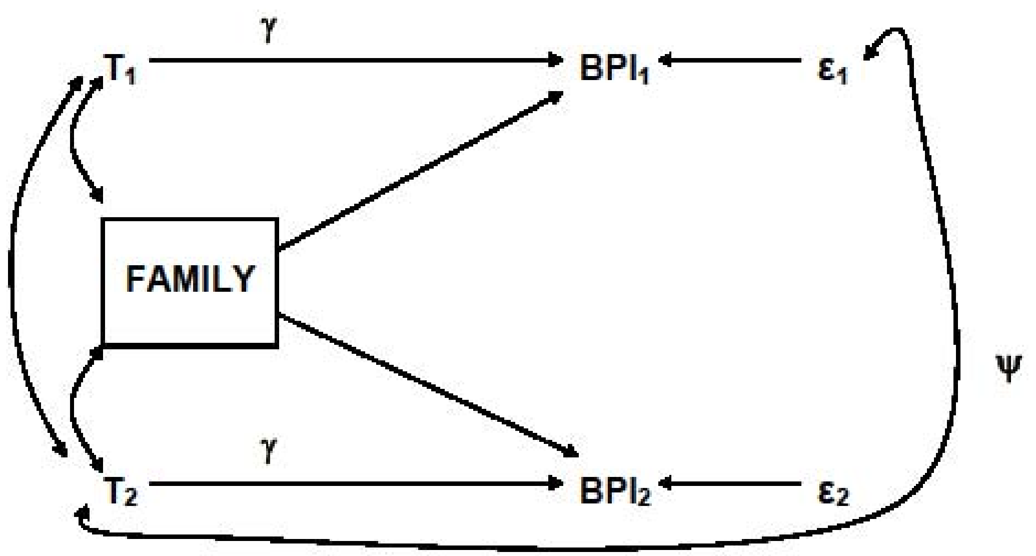


Figure 1. Path Diagram of a Simple Fixed-Effects Model Incorporating Dynamic Decision-making



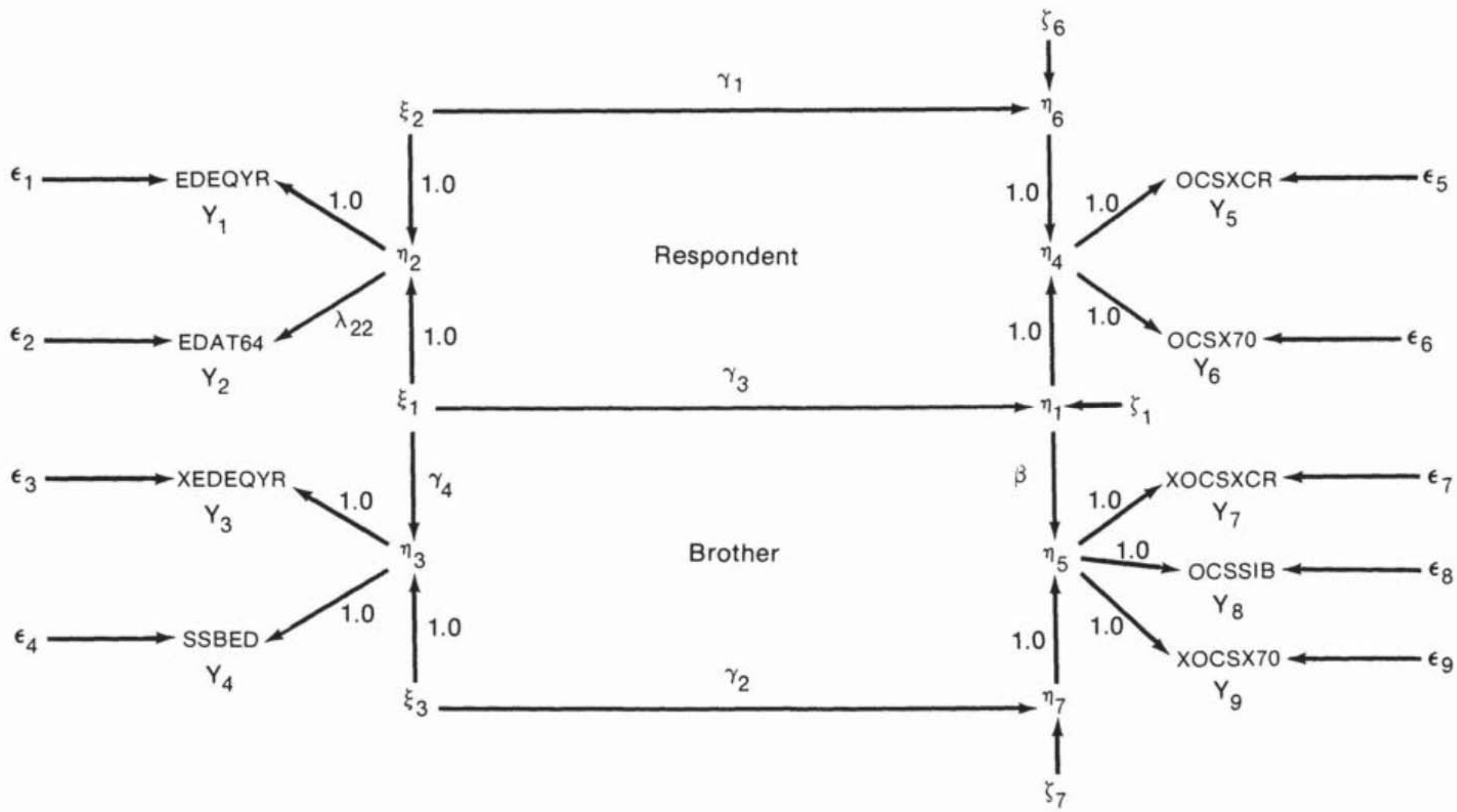


FIG. 2.—A model of sibling resemblance in educational attainment and occupational status with errors in variables and latent family factors.