Klein, S. P. and Freedman, D. A. (1993), "Ecological regression in voting rights cases" Chance, 6, 38-43.

Additional Resources
Aggregation bias, Ecological fallacy.
D.A. Freedman. "The ecological fallacy." In the Encyclopedia of Social Science Research Methods. Sage Publications (2004) Vol. 1 p. 293. M. Lewis-Beck, A. Bryman, and T. F. Liao, eds
A Rule for Inferring Individual-Level Relationships from Aggregate Data, Glenn Firebaugh *American Sociological Review* Vol. 43, No. 4 (Aug., 1978), pp. 557-572 JStor URL
American Journal of Epidemiology Vol. 139, No. 8: 747-760 Invited Commentary: Ecologic Studies -- Biases, Misconceptions, and Counterexamples S Greenland, J Robins
The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology J. Michael Oakes *Social Science and Medicine* 58 (2004) 1929–1952
R-package eiPack: R x C Ecological Inference and Higher-Dimension Data Management. R News Oct 2007
Educational multilevel data.
The Analysis of Multilevel Data in Educational Research and Evaluation Leigh Burstein *Review of Research in Education*, Vol. 8. (1980), pp. 158-233. Jstor link
Methodological Advances in Analyzing the Effects of Schools and Classrooms on Student Learning, Stephen W. Raudenbush; Anthony S. Bryk Review of Research in Education, Vol. 15. (1988 - 1989), pp. 423-475. Jstor link
Analyzing Multilevel Data in the Presence of Heterogeneous within-Class Regressions Leigh Burstein; Robert L. Linn; Frank J. Capell *Journal of Educational Statistics*, Vol. 3, No. 4. (Winter, 1978), pp. 347-383. Jstor link
examples from analyses of voting data.
Bias in ecological regression Stephen Ansolabehere and Douglas Rivers
David A. Freedman et al., "Ecological Regression and Voting Rights," *Evaluation Review* 1991, pp. 673-711,
D.A. Freedman, S.P. Klein, M. Ostland, and M.R. Roberts. "Review of 'A Solution to the Ecological Inference Problem.' " *Journal of the American Statistical Association*, vol. 93 (1998) pp. 1518-22; with discussion, vol. 94 (1999) pp. 352–57.
Multilevel models.
Using SAS PROC mixed:
Judith Singer HLM/PROC Mixed papers: Multilevel Modelling Newsletter ; JEBS1998 Using SAS PROC MIXED to Fit Multilevel Models, Jstor
HLM - Hierarchical Linear and Nonlinear Modeling (HLM): descriptions and student edition HLM6
Freedman, D. A. (census adjustments). Hierarchical Linear Regression
Using R: lme4 (lmer and nlme) and mlmRev. John Fox lme tutorial
Doug Bates draft book (Feb 2010) Doug Bates SASmixed package
Fitting linear mixed models in R Using the lme4 package Douglas Bates (pp.27-30)
London exam data example in Examples from Multilevel Software Comparative Reviews Douglas Bates
Regression diagnostics for lmer models. Package influence.ME
mlmRev data examples. Also, Tennessee's Student Teacher Achievement Ratio (STAR) from Creating an R data set from STAR Douglas Bates
STATA does it also
lmer for SAS PROC MIXED Users Douglas Bates Department of Statistics University of Wisconsin Madison

2. Reciprocal Causal Effects and non-recursive models in Observational Studies

Lecture topics
1. Cross-sectional Data: Simultaneous equations (2SLS, IV in butter, peer aspirations, ed and fertility, Freedman), nonrecursive models
   Simultaneous equations handouts Duncan et al ascii
2. Reciprocal effects and non-recursive models in longitudinal data.
   Empirical research on reciprocal effects, including cross-lagged correlation. clc slides

Primary Readings
An (old) review of reciprocal effects. Rogosa, D. R. (1985). Analysis of reciprocal effects. In International Encyclopedia of Education, T. Husen and N. Postlethwaite, Eds. London: Pergamon Press, 4221-4225. (reprinted in Educational Research,Methodology & Measurement: An international handbook, J. P. Keeves Ed. Oxford: Pergamon Press, 1988.)

Reciprocal Effects Examples
Michelob ULTRA® Super Bowl LV Spot Online. Are You Happy Because You Win? Or Do You Win Because You're Happy?
Screen time rots kids minds.
Fox17 Nashville: Increased screen time in young children associated with developmental delays
Publication: Association Between Screen Time and Children's Performance on a Developmental Screening Test JAMA Pediatr. Published online January 28, 2019. doi:10.1001/jamapediatrics.2018.5056
Internet use and depression
Study links excessive internet use to depression Publication: The Relationship between Excessive Internet Use and Depression: A Questionnaire-Based Study of 1,319 Young People and Adults. Catriona M. Morrison, Helen Gore *Psychopathology* 2010;43:121-126 . available from Lane e-journals
Peer Influences
Peer Influences on Aspirations: A Reinterpretation Otis Dudley Duncan, Archibald O. Haller, Alejandro Portes American Journal of Sociology, Vol. 74, No. 2 (Sep., 1968), pp. 119-137 Jstor
Education and Fertility
Rindfus example (Freedman Chap 8; paper reprinted in Freedman text). Education and Fertility: Implications for the Roles Women Occupy Ronald R. Rindfuss; Larry Bumpass; Craig St. John American Sociological Review, Vol. 45, No. 3. (Jun., 1980), pp. 431-447. from Jstor
Longitudinal Data: original TV Violence and Agression
Eron LD, Huesmann LR, Lefkowitz MM, Walder LO. Does television violence cause aggression? Am Psychol. 1972;27:253–63. PubMed
Money Supply
Granger Causality. Nobel 2003. Complete Granger
Relationships--and the Lack Thereof--Between Economic Time Series, with Special Reference to Money and Interest Rates. David A. Pierce *Journal of the American Statistical Association*, Vol. 72, No. 357. (Mar., 1977), pp. 11-26. Jstor

Additional Resources
Reciprocal effects: Rogosa, D. R. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, 88, 245-258. APA site version
Structural Equation Modeling With the sem Package in R John Fox STRUCTURAL EQUATION MODELING,13(3),465–486 Jox Fox home page

Weeks 6 [redacted]

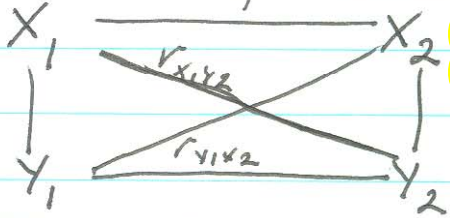
Stat 209 [redacted]

Reciprocal Effects, Simultaneous Equations

CLC (cross-lagged correlation)

compare $r_{X_1 Y_2}$ to $r_{Y_1 X_2}$

causal predominance to the larger (sig)



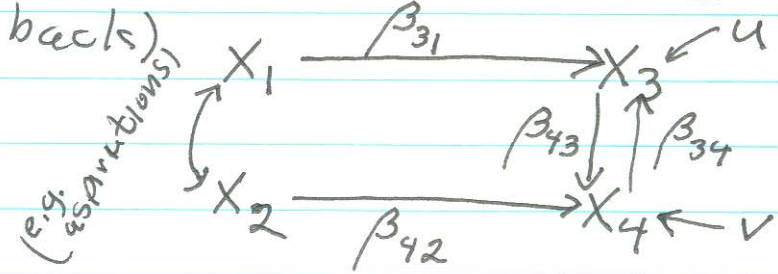
Longitudinal data

Non-Recursive (feedback)

$$X_3 = \beta_{31} X_1 + \beta_{34} X_4 + u$$

$$X_4 = \beta_{42} X_2 + \beta_{43} X_3 + v$$

$E(X_1) = 0 \quad E(u) = E(v) = 0$



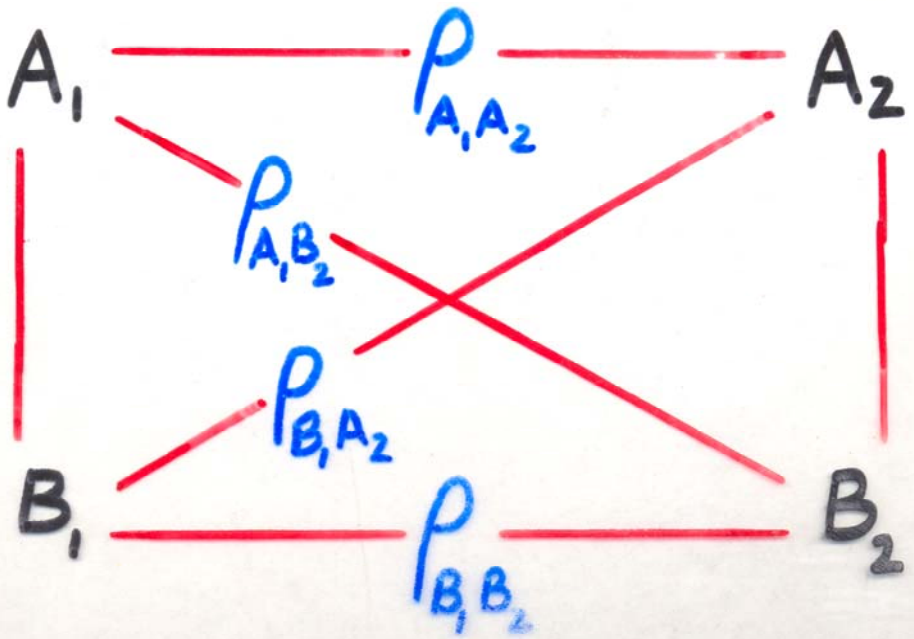
Reciprocal Effects

Does A cause B or does B cause A?

TV violence and aggression?

Personality characteristics and drug use?

Self esteem and achievement?



Study links excessive internet use to depression

02 February 2010 @ 07:56 pm ET

Next Tech Article

LONDON - People who spend a lot of time surfing the internet are more likely to show signs of depression, British scientists said on Wednesday.

People use computers at an internet cafe in Taiyuan, Shanxi province, November 13, 2009. REUSTERS/Stringer (REUTERS)

1 of 1 Full Size

But it is not clear whether the internet causes depression or whether depressed people are drawn to it.

Psychologists from Leeds University found what they said was "striking" evidence that some avid net users develop compulsive internet habits in which they replace real-life social interaction with online chat rooms and social networking sites.

"This study reinforces the public speculation that over-engaging in websites that serve to replace normal social function might be linked to psychological disorders like depression and addiction," the study's lead author, Catriona Morrison, wrote in the journal Psychopathology.

"This type of addictive surfing can have a serious impact on mental health."

In the first large-scale study of Western young people to look at this issue, the researchers analyzed internet use and depression levels of 1,319 Britons aged between 16 and 51.

Of these, 1.2 percent were "internet addicted", they concluded.

These "internet addicts" spent proportionately more time browsing sexually gratifying websites, online gaming sites and online communities, Morrison said. They also had a higher incidence of moderate to severe depression than normal users.

"Excessive internet use is associated with depression, but what we don't know is which comes first -- are depressed people drawn to the internet or does the internet cause depression?," Morrison said.

"What is clear is that for a small subset of people, excessive use of the internet could be a warning signal for depressive tendencies."

Morrison noted that while the 1.2 percent figure for those classed as "addicts" was small, it was larger than the incidence of gambling in Britain, which is around 0.6 percent. (Reporting by Kate Kelland. Editing by Paul Casciato)

Tech Blog

Apple's iPad real cost is $270

LIVE BLOG: Apple Announcement today at 10 a.m. PT / 1 p.m. ET

NYTimes announces pay model for website

Read More »

More Tech

Iran students hold anti-government protest: website

Students demonstrated at Tehran University against the government of President Mahmoud Ahmadinejad on Monday, a reformist website reported, more than thr...

Opera urges EU regulators not to rush Microsoft case

Norwegian browser maker Opera urged European Union antitrust regulators on Monday not to rush to close its antitrust case against Microsoft before ensuri...

Apple Breaks Through 2 Billion Apps Mark

Apple announced Monday that over 2 billion apps have been downloaded from its App Store since the e-commerce site's launch in July 2008, and the pace "co...

RSS Feed Link

Table 3. Comparison of proportion of time spent on different Internet activities for IA and NA groups

	NA group	IA group	Summary statistics		
			MS	F	p
Sexually gratifying sites	1.36 (0.169) 0.99–1.72	3.33 (0.513) 2.23–4.43	28.28	(1, 27) = 12.611	0.001
Games	2.07 (0.474) 1.05–3.10	4.00 (0.412) 3.13–4.87	29.290	(1, 30) = 9.456	<0.005
Chat	2.13 (0.487) 1.09–3.18	4.13 (0.496) 3.07–5.20	30.000	(1, 28) = 8.279	<0.01
Browsing	2.56 (0.353) 1.81–3.32	3.72 (0.360) 2.96–4.48	11.393	(1, 32) = 5.242	<0.05
Community	2.94 (0.431) 2.03–3.85	4.24 (0.442) 3.30–5.17	14.568	(1, 33) = 4.37	<0.05
E-mail	2.27 (0.419) 1.37–3.17	1.78 (0.275) 1.20–2.36	1.956	(1, 31) = 1.01	>0.1
Research	2.5 (0.303) 1.85–3.15	2.13 (0.424) 1.22–3.04	1.041	(1, 29) = 0.505	>0.1
Gambling	1.0 (0.000) 1.00–1.00	1.45 (0.455) 0.44–2.47	1.082	(1, 19) = 0.905	>0.1
eBay/shopping	1.12 (0.081) 0.95–1.29	1.07 (0.067) 0.92–1.21	0.021	(1, 30) = 0.23	>0.1
Other	1.67 (0.333) 0.93–2.40	2.50 (0.562) 1.29–3.71	4.487	(1, 24) = 1.492	>0.1

Values in the NA and IA group columns are means (SEs in parentheses), on a scale of 1 = rarely/never to 6 = very frequently. Confidence intervals are listed beneath means and SEs.

table 3. These data back up the evidence from the correlation matrix, showing that the IA group spent significantly more time on sexually gratifying sites, gaming sites, chat sites, browsing and community sites. The activities for which the NA group had a preference did not differ significantly, indicating that the NA group have a more even spread of activities.

Discussion

In summary, we found a clear link between IA and depression, such that those whom we classed as addicted were significantly more depressed than those in the NA group. Hence we have identified a statistically significant relationship between IA and depression. What is not clear from these data is which comes first: are depressed people drawn to the Internet, or does excessive Internet use make one more prone to depression? This needs further work in the future, but it is clear that, for a small subgroup of the population, excessive use of the Internet is a warning signal of depressive tendencies. However, in line with previous studies, this subgroup represents <2% of the population. This is the figure typically reported in the literature, and it is higher than the incidence of gambling in the UK, which stands at around 0.6% [16].

When considering the functions of the Internet, the important point to note is that there was a significant difference between the groups in terms of sexually gratifying websites, online games and chat/community sites, such that the IA group engaged significantly more in these sites than did the NA group. This accords with recent evidence suggesting that those prone to dependence on the Internet are drawn to sites that involve these 3 types of activities [2]. This feeds the public speculation that overengagement in websites that serve/replace a so-

Weeks 6 [redacted]

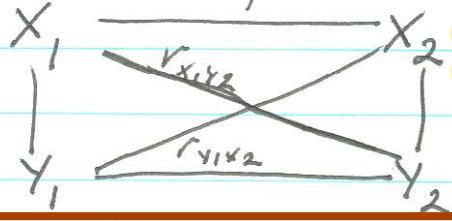
Stat 209 [redacted]

Reciprocal Effects, Simultaneous Equations

CLC (cross-lagged correlation)

compare $r_{X_1 Y_2}$ to $r_{Y_1 X_2}$

causal predominance to the larger (sig)

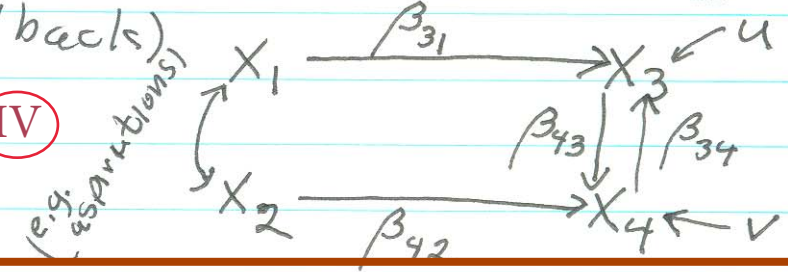


Longitudinal data

Non-Recursive (feedback)

$X_3 = \beta_{31} X_1 + \beta_{34} X_4 + u$ (IV)

$X_4 = \beta_{42} X_2 + \beta_{43} X_3 + v$



$E(X_1) = 0 \quad E(u) = E(v) = 0$

Weeks 6 [redacted]

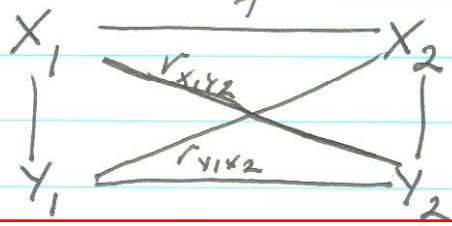
Stat 209 [redacted]

Reciprocal Effects, Simultaneous Equations

CLC (cross-lagged correlation)

compare $r_{X_1 Y_2}$ to $r_{Y_1 X_2}$ (sig)

causal predominance to the larger



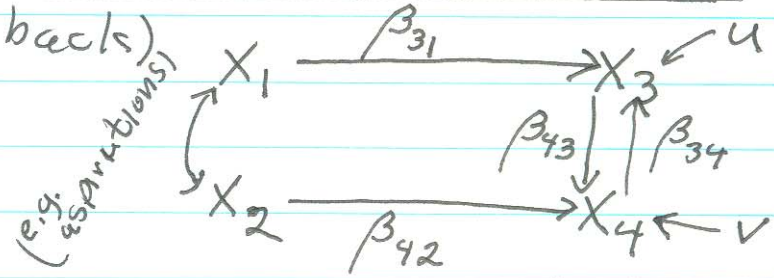
Non-Recursive (feedback)

$X_3 = \beta_{31} X_1 + \beta_{34} X_4 + u$

$X_4 = \beta_{42} X_2 + \beta_{43} X_3 + v$

$E(X_i) = 0 \quad E(u) = E(v) = 0$

(just ident $\sigma_{uv} = 0$)



X_1, X_2 exogenous $\quad X_3, X_4$ endogenous

u, v uncorr w/ exogenous

OLS estimates

Simultaneity Bias

$$E(\hat{\beta}_{31}^{OLS}) = \beta_{31} - \frac{\sigma_{14}\sigma_{4u}}{\sigma_{11}\sigma_{44} - \sigma_{14}^2}$$

$$E(\hat{\beta}_{34}^{OLS}) = \beta_{34} + \frac{\sigma_{11}\sigma_{4u}}{\sigma_{11}\sigma_{44} - \sigma_{14}^2}$$

$$E(\hat{\beta}_{42}^{OLS}) = \beta_{42} - \frac{\sigma_{23}\sigma_{3v}}{\sigma_{22}\sigma_{33} - \sigma_{23}^2}$$

$$E(\hat{\beta}_{43}^{OLS}) = \beta_{43} + \frac{\sigma_{22}\sigma_{3v}}{\sigma_{22}\sigma_{33} - \sigma_{23}^2}$$

By endogeneity $\sigma_{34} = \beta_{34}\sigma_{44} + \sigma_{4u}$

$\sigma_{44} = \beta_{43}\sigma_{34} + \sigma_{uv}$, $\sigma_{3v} = \beta_{34}\sigma_{4v} + \sigma_{uv}$

$\sigma_{4v} = \beta_{43}\sigma_{3v} + \sigma_{vv}$

IV estimates (consistent)

$$\hat{\beta}_{31}^{IV} = \frac{s_{13}s_{24} - s_{14}s_{23}}{s_{11}s_{24} - s_{12}s_{14}}$$

(IV sub x_2 for x_4)

whereas $\hat{\beta}_{31.4}^{OLS} = \dfrac{s_{13}s_{44} - s_{14}s_{34}}{s_{11}s_{44} - s_{14}^2}$

$$\hat{\beta}_{34}^{IV} = \frac{s_{11}s_{23} - s_{12}s_{13}}{s_{11}s_{24} - s_{12}s_{14}} \qquad \hat{\beta}_{42}^{IV} = \frac{s_{14}s_{23} - s_{13}s_{24}}{s_{12}s_{23} - s_{13}s_{22}} \qquad \hat{\beta}_{43}^{IV} = \frac{s_{12}s_{24} - s_{14}s_{22}}{s_{12}s_{23} - s_{13}s_{22}}$$

Simultaneous Eq's

Woolridge p.562 linked in Lab session

Lab 3 (Task 3) PSID data

$$hours = \alpha_1 \log Wage + \beta_{10} + \beta_{11} educ + \beta_{12} age + \beta_{13} kids5 + \beta_{14} nwife + u$$

$$\log Wage = \alpha_2 hours + \beta_{20} + \beta_{21} educ + \beta_{22} exper + \beta_{23} exper^2 + v$$

Duncan - Occupational aspiration; Rindflus - Education, fertility (Freedman)

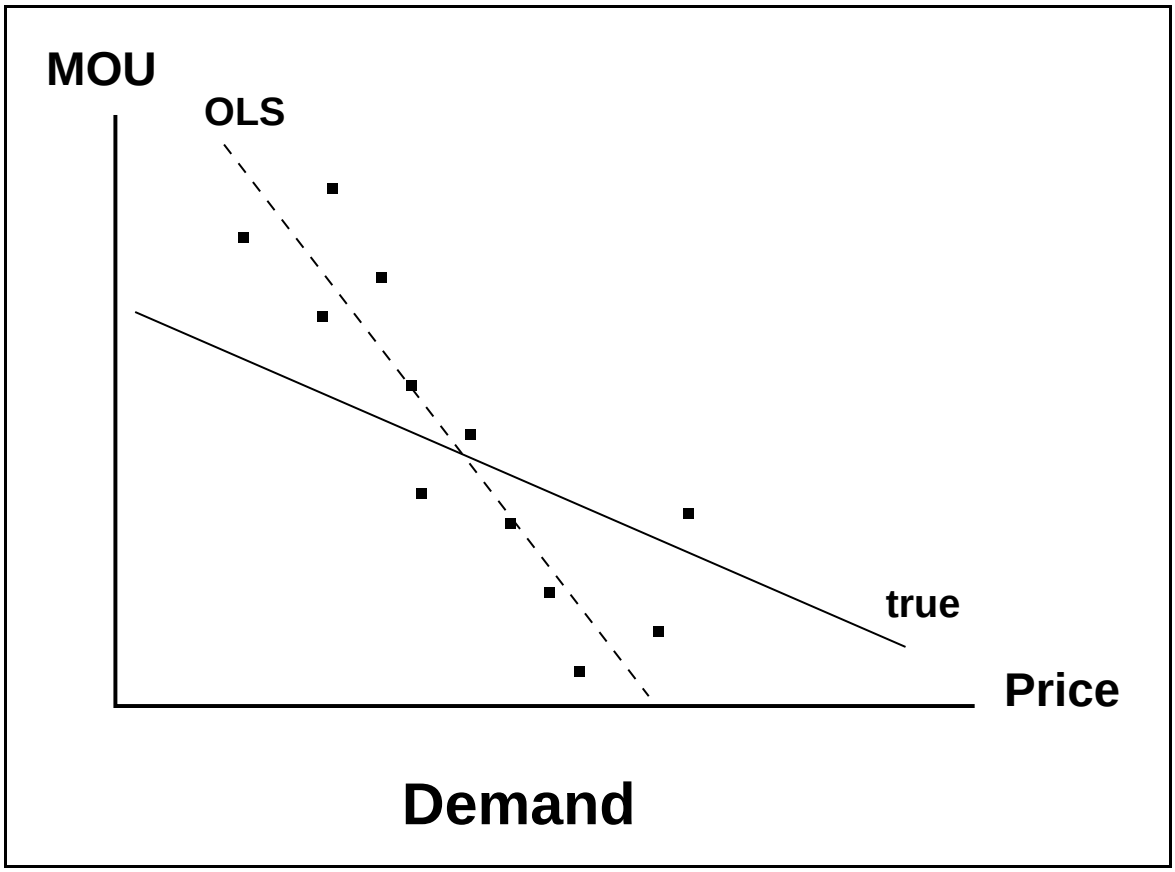
Example 2

Price elasticity of telecommunications demand:

$$MOU = \alpha + \beta(price) + \theta(Nemply) + \varepsilon$$

$$Price = \lambda + \varphi(MOU) + \eta(region\ dum.) + \mu$$

ε and μ are correlated.



Problem: Because μ is correlated with ε, and μ affects price, price is correlated with ε.

Linked in Lab3

In a nutshell, this illustrates the identification problem in simultaneous equations models, which we will discuss more generally in Section 16.3.

The most convincing examples of SEMs have the same flavor as supply and demand examples. Each equation should have a behavioral, ceteris paribus interpretation on its own. Because we only observe equilibrium outcomes, specifying an SEM requires us to ask such counterfactual questions as: How much labor *would* workers provide if the wage were different from its equilibrium value? Example 16.1 provides another illustration of an SEM where each equation has a ceteris paribus interpretation.

Reciprocal Effects Simultaneous Eqs

EXAMPLE 16.1

(Murder Rates and Size of the Police Force)

Cities often want to determine how much additional law enforcement will decrease their murder rates. A simple cross-sectional model to address this question is

$$murdpc = \alpha_1 polpc + \beta_{10} + \beta_{11} incpc + u_1, \tag{16.6}$$

where *murdpc* is murders per capita, *polpc* is number of police officers per capita, and *incpc* is income per capita. (Henceforth, we do not include an *i* subscript.) We take income per capita as exogenous in this equation. In practice, we would include other factors, such as age and gender distributions, education levels, perhaps geographic variables, and variables that measure severity of punishment. To fix ideas, we consider equation (16.6).

The question we hope to answer is: If a city exogenously increases its police force, will that increase, on average, lower the murder rate? If we could exogenously choose police force sizes for a random sample of cities, we could estimate (16.6) by OLS. Certainly, we cannot run such an experiment. But can we think of police force size as being exogenously determined, anyway? Probably not. A city's spending on law enforcement is at least partly determined by its expected murder rate. To reflect this, we postulate a second relationship:

$$polpc = \alpha_2 murdpc + \beta_{20} + other\ factors. \tag{16.7}$$

We expect that $\alpha_2 > 0$: other factors being equal, cities with higher (expected) murder rates will have more police officers per capita. Once we specify the other factors in (16.7), we have a two-equation simultaneous equations model. We are really only interested in equation (16.6), but, as we will see in Section 16.3, we need to know precisely how the second equation is specified in order to estimate the first.

An important point is that (16.7) describes behavior by city officials, while (16.6) describes the actions of potential murderers. This gives each equation a clear ceteris paribus interpretation, which makes equations (16.6) and (16.7) an appropriate simultaneous equations model.

We next give an example of an inappropriate use of SEMs.

non-recursive models, feedback

Cross-sectional Data

16.2 SIMULTANEITY BIAS IN OLS

It is useful to see, in a simple model, that an explanatory variable that is determined simultaneously with the dependent variable is generally correlated with the error term, which leads to bias and inconsistency in OLS. We consider the two-equation structural model

are y_2 , u_1 correlated? bias, see handout

$$y_1 = \alpha_1 y_2 + \beta_1 z_1 + u_1 \tag{16.10}$$

$$y_2 = \alpha_2 y_1 + \beta_2 z_2 + u_2 \tag{16.11}$$

and focus on estimating the first equation. The variables z_1 and z_2 are exogenous, so that each is uncorrelated with u_1 and u_2. For simplicity, we suppress the intercept in each equation.

To show that y_2 is generally correlated with u_1, we solve the two equations for y_2 in terms of the exogenous variables and the error term. If we plug the right-hand side of (16.10) in for y_1 in (16.11), we get

$$y_2 = \alpha_2(\alpha_1 y_2 + \beta_1 z_1 + u_1) + \beta_2 z_2 + u_2$$

or

$$(1 - \alpha_2\alpha_1)y_2 = \alpha_2\beta_1 z_1 + \beta_2 z_2 + \alpha_2 u_1 + u_2. \tag{16.12}$$

Now, we must make an assumption about the parameters in order to solve for y_2:

$$\alpha_2\alpha_1 \neq 1. \tag{16.13}$$

Whether this assumption is restrictive depends on the application. In Example 16.1, we think that $\alpha_1 \leq 0$ and $\alpha_2 \geq 0$, which implies $\alpha_1\alpha_2 \leq 0$; therefore, (16.13) is very reasonable for Example 16.1.

Provided condition (16.13) holds, we can divide (16.12) by $(1 - \alpha_2\alpha_1)$ and write y_2 as

$$y_2 = \pi_{21} z_1 + \pi_{22} z_2 + v_2, \tag{16.14}$$

where $\pi_{21} = \alpha_2\beta_1/(1 - \alpha_2\alpha_1)$, $\pi_{22} = \beta_2/(1 - \alpha_2\alpha_1)$, and $v_2 = (\alpha_2 u_1 + u_2)/(1 - \alpha_2\alpha_1)$. Equation (16.14), which expresses y_2 in terms of the exogenous variables and the error terms, is the **reduced form** for y_2, a concept we introduced in Chapter 15 in the context of instrumental variables estimation. The parameters π_{21} and π_{22} are called **reduced form parameters**: notice how they are nonlinear functions of the **structural parameters**, which appear in the structural equations, (16.10) and (16.11).

The **reduced form error**, v_2, is a linear function of the structural error terms, u_1 and u_2. Because u_1 and u_2 are each uncorrelated with z_1 and z_2, v_2 is also uncorrelated with z_1 and z_2. Therefore, we can consistently estimate π_{21} and π_{22} by OLS, something that is used for two stage least squares estimation (which we return to in the next section). In addition, the reduced form parameters are sometimes of direct interest, although we are focusing here on estimating equation (16.10).

TSLS multiple instruments

$$y_1 = \beta_0 + \beta_1 y_2 + \beta_2 z_1 + u_1$$

y_2 correl u_1 endogenous e.g. simultaneous eqs

instruments z_2, z_3 use jointly

$$y_2 = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + \pi_3 z_3 + v_2$$

expect is y_2^* part of y_2 uncorrelated with u_1

$E(y_2 | z_i)$ as all z_i uncorr w u_1 (assumption)

estimate y_2^* by $\hat{y}_2$ fit y_2 on $z_1\ z_2\ z_3$

use $\hat{y}_2$ as IV for y_2

estimate β_1 from y_1 on $\hat{y}_2\ z_1$ regressions

hence, TSLS

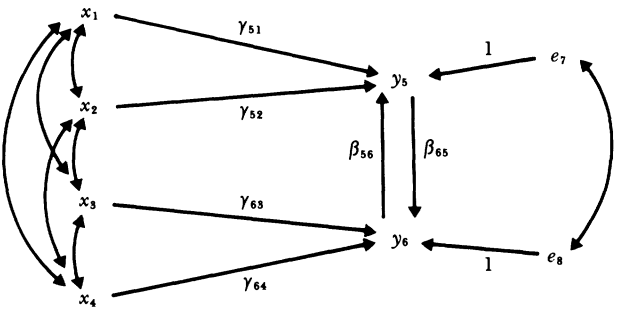
cf Lab 3

simultaneous eqs, tsls, for tues

Fox 1979 Soc Meth (link)

randout

Figure 1. Overidentified nonrecursive model.



x_1: Respondent's intelligence
x_2: Respondent's family SES
x_3: Friend's family SES
x_4: Friend's intelligence
y_5: Respondent's occupational aspiration
y_6: Friend's occupational aspiration

SOURCE: Duncan and others (1971). The symbols used here are different from those in the original source.

Two points are worth mentioning. First, the signs of the coefficients attached to endogenous and exogenous independent variables in an equation have been reversed, since in going from the path diagram and scalar equation representations of the model to the matrix representation, we have shifted these variables from the right to the left-hand side of the equation. Second, I have omitted constant (that is, intercept) terms from each structural equation and hence we shall stipulate that all variables be measured as deviations from their means. In a later section, I shall see how a constant term may be recovered for each equation.

Since the $\mathbf{B}$ matrix is nonsingular, it is possible to solve the structural equations for the endogenous variables. This operation, shown below, produces the "reduced form" of the model, expressing the endogenous variables in terms of the exogenous variables and disturbances that, together, comprise the ultimate inputs to the system.

$$\mathbf{Y} = -\mathbf{X\Gamma B}^{-1} + \mathbf{EB}^{-1} = \mathbf{X\Pi} + \mathbf{V} \qquad (4)$$

where $\mathbf{\Pi} = -\mathbf{\Gamma B}^{-1}$ and $\mathbf{V} = \mathbf{EB}^{-1}$. Then $\mathbf{\Pi}$ is the matrix of

$$
\begin{aligned}
y_5 &= \underset{(0.104)}{0.403}y_6 + \underset{(0.053)}{0.272}x_1 + \underset{(0.054)}{0.151}x_2 + 0.841\hat{e}'_7 \\
y_6 &= \underset{(0.125)}{0.341}y_5 + \underset{(0.054)}{0.157}x_3 + \underset{(0.055)}{0.352}x_4 + 0.805\hat{e}'_8
\end{aligned}
\tag{27}
$$

These estimates agree with those given in the original article, although standard errors are not reported there. The correlation between the disturbances, $r_{e'_7e'_8} = -0.476$, also agrees with the value reported by Duncan and colleagues. The overidentification F statistics for the two equations are 2.51 and 1.77, both with 1 and 324 degrees of freedom (and neither statistically significant).

TSLS multiple instruments

$$y_1 = \beta_0 + \beta_1 y_2 + \beta_2 z_1 + u_1$$

y_2 correl u_1 endogenous e.g. simultaneous eqs

instruments z_2, z_3 use jointly

$$y_2 = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + \pi_3 z_3 + v_2$$

expect is y_2^* part of y_2 uncorrelated with u_1

$E(y_2 | z_i)$ as all z_i uncorr w u_1 (assumption)

estimate y_2^* by $\hat{y}_2$ fit y_2 on $z_1\ z_2\ z_3$

use $\hat{y}_2$ as IV for y_2

estimate β_1 from y_1 on $\hat{y}_2\ z_1$ regressions

hence, TSLS

cf Lab 3

```
> # Now to Duncan Haller Portes 1968 (Peer influences on Aspirations)
> # path diagram and data from Fox Soc Meth 1979
>  dunccor = matrix(nrow = 6, ncol = 6, c(1,.222,.1861,.3355,.4105,.2598,.222,1,.2707,.2302,.3240,.27
+  .1861,.2707,1,.2950,.293,.3607,.3355,.2302,.2950,1..2995,.5007,.4105,.3240,.2930,.2995,1,.4216,
+  .2598,.2786,.3607,.5007,.4216,1))
> dunccor
       [,1]   [,2]   [,3]   [,4]   [,5]   [,6]
[1,] 1.0000 0.2220 0.1861 0.3355 0.4105 0.2598
[2,] 0.2220 1.0000 0.2707 0.2302 0.3240 0.2786
[3,] 0.1861 0.2707 1.0000 0.2950 0.2930 0.3607
[4,] 0.3355 0.2302 0.2950 1.0000 0.2995 0.5007
[5,] 0.4105 0.3240 0.2930 0.2995 1.0000 0.4216
[6,] 0.2598 0.2786 0.3607 0.5007 0.4216 1.0000
```

x_1: Respondent's intelligence
x_2: Respondent's family SES
x_3: Friend's family SES
x_4: Friend's intelligence
y_5: Respondent's occupational aspiration
y_6: Friend's occupational aspiration

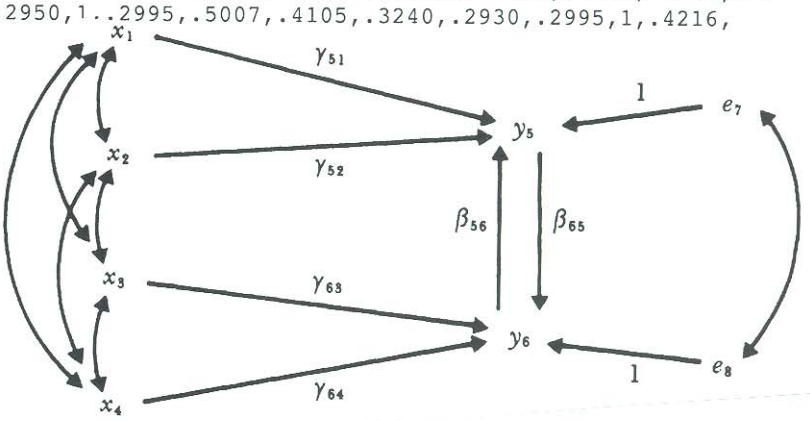


Figure 1. Overidentified nonrecursive model.

```
> library(MASS)
> duncdatemp = mvrnorm(329, c(0,0,0,0,0,0),dunccor, empirical = TRUE)
> cor(duncdatemp) # matches above
```

```
> focreg = tsls(duncdatemp[,6] ~ duncdatemp[,3] + duncdatemp[,4] + duncdatemp[,5],
+ ~ duncdatemp[,1] +duncdatemp[,2] + duncdatemp[,3] + duncdatemp[,4])
> summary(focreg)
 2SLS Estimates
Model Formula: duncdatemp[, 6] ~ duncdatemp[, 3] + duncdatemp[, 4] + duncdatemp[,5]
Instruments: ~duncdatemp[, 1] + duncdatemp[, 2] + duncdatemp[, 3] + duncdatemp[,4]
                  Estimate Std. Error    t value  Pr(>|t|)
(Intercept)     -1.896e-17    0.04457 -4.254e-16 1.000e+00
duncdatemp[, 3]  1.567e-01    0.05445  2.877e+00 4.278e-03
duncdatemp[, 4]  3.521e-01    0.05505  6.396e+00 5.554e-10
duncdatemp[, 5]  3.419e-01    0.12478  2.740e+00 6.484e-03
Residual standard error: 0.8084 on 325 degrees of freedom

> #got to have as many instruments as predictors
> rocreg = tsls(duncdatemp[,5] ~ duncdatemp[,1] + duncdatemp[,2] + duncdatemp[,6],
+ ~ duncdatemp[,1] + duncdatemp[,2] + duncdatemp[,3] + duncdatemp[,4])
> summary(rocreg)
 2SLS Estimates
Model Formula: duncdatemp[, 5] ~ duncdatemp[, 1] + duncdatemp[, 2] + duncdatemp[,6]
Instruments: ~duncdatemp[, 1] + duncdatemp[, 2] + duncdatemp[, 3] + duncdatemp[,4]
                  Estimate Std. Error    t value  Pr(>|t|)
(Intercept)     -1.859e-17    0.04658 -3.991e-16 1.000e+00
duncdatemp[, 1]  2.721e-01    0.05255  5.179e+00 3.923e-07
duncdatemp[, 2]  1.512e-01    0.05364  2.819e+00 5.113e-03
duncdatemp[, 6]  4.034e-01    0.10431  3.867e+00 1.330e-04
Residual standard error: 0.8449 on 325 degrees of freedom

> # both 2SLS results match Fox Soc Meth p.145 results
OLS Comparisons
lm(formula = duncdatemp[, 5] ~ duncdatemp[, 1] + duncdatemp[,2] + duncdatemp[, 6])
Coefficients:     Estimate Std. Error    t value Pr(>|t|)
(Intercept)     -2.032e-17  4.625e-02 -4.39e-16 1.000000
duncdatemp[, 1]  2.945e-01  4.860e-02      6.059 3.78e-09 ***
duncdatemp[, 2]  1.762e-01  4.887e-02      3.605 0.000361 ***
duncdatemp[, 6]  2.960e-01  4.934e-02      5.999 5.28e-09 ***
Residual standard error: 0.8389 on 325 degrees of freedom
lm(formula = duncdatemp[, 6] ~ duncdatemp[, 3] + duncdatemp[,4] + duncdatemp[, 5])
Coefficients:     Estimate Std. Error    t value Pr(>|t|)
(Intercept)     -1.308e-17  4.436e-02 -2.95e-16 1.000000
duncdatemp[, 3]  1.752e-01  4.772e-02      3.672 0.000281 ***
duncdatemp[, 4]  3.714e-01  4.782e-02      7.767 1.06e-13 ***
duncdatemp[, 5]  2.590e-01  4.779e-02      5.420 1.16e-07 ***
Residual standard error: 0.8047 on 325 degrees of freedom
```

OLS
5,6 weights a little different

C1 Freedman, case study Rindfus Ed + Fertility ASR, Stat 209 2/14/08

```
> rindcor = matrix(nrow = 11, ncol = 11,
+ c( 1.000,-0.144,-0.244,-0.323,-0.129,-0.056 , 0.053 ,-0.043 , 0.037 , 0.370 , 0.186,
+ -0.144, 1.000 , 0.156, 0.088, 0.315, 0.150,-0.152, 0.030, 0.035,-0.222,-0.189,
+ -0.244,   0.156,   1.000, 0.274, 0.150,-0.039, 0.014, 0.028, 0.002,-0.328,-0.115,
+ -0.323,   0.088,   0.274,   1.000, 0.218,-0.030,-0.149,-0.060,-0.032,-0.185,-0.118,
+ -0.129, 0.315, 0.150, 0.218, 1.000, 0.071,-0.292,-0.011,-0.027,-0.211,-0.177,
+ -0.056, 0.150,  -0.039,  -0.030, 0.071, 1.000,-0.052, 0.067, 0.018,-0.157, 0.111,
+  0.053,  -0.152,   0.014,  -0.149,  -0.292,  -0.052,   1.000,-0.010,-0.002,-0.012, 0.098,
+ -0.043,   0.030,   0.028,  -0.060,  -0.011,   0.067,  -0.010,   1.000, 0.009,-0.171,-0.122,
+ 0.037,   0.035,   0.002,  -0.032,  -0.027,   0.018,  -0.002,   0.009,   1.000,  0.038, 0.216,
+  0.370,  -0.222,  -0.328,  -0.185,  -0.211,  -0.157,  -0.012,  -0.171,   0.038,   1.000, 0.380,
+ 0.186,  -0.189,  -0.115,  -0.118,  -0.177,   0.111,   0.098,  -0.122,   0.216,   0.380, 1.0))
> rindcor
        [,1]   [,2]   [,3]   [,4]   [,5]   [,6]   [,7]   [,8]   [,9]  [,10]  [,11]
 [1,]  1.000 -0.144 -0.244 -0.323 -0.129 -0.056  0.053 -0.043  0.037  0.370  0.186
 [2,] -0.144  1.000  0.156  0.088  0.315  0.150 -0.152  0.030  0.035 -0.222 -0.189
 [3,] -0.244  0.156  1.000  0.274  0.150 -0.039  0.014  0.028  0.002 -0.328 -0.115
 [4,] -0.323  0.088  0.274  1.000  0.218 -0.030 -0.149 -0.060 -0.032 -0.185 -0.118
 [5,] -0.129  0.315  0.150  0.218  1.000  0.071 -0.292 -0.011 -0.027 -0.211 -0.177
 [6,] -0.056  0.150 -0.039 -0.030  0.071  1.000 -0.052  0.067  0.018 -0.157  0.111
 [7,]  0.053 -0.152  0.014 -0.149 -0.292 -0.052  1.000 -0.010 -0.002 -0.012  0.098
 [8,] -0.043  0.030  0.028 -0.060 -0.011  0.067 -0.010  1.000  0.009 -0.171 -0.122
 [9,]  0.037  0.035  0.002 -0.032 -0.027  0.018 -0.002  0.009  1.000  0.038  0.216
[10,]  0.370 -0.222 -0.328 -0.185 -0.211 -0.157 -0.012 -0.171  0.038  1.000  0.380
[11,]  0.186 -0.189 -0.115 -0.118 -0.177  0.111  0.098 -0.122  0.216  0.380  1.000
> rinddat = mvrnorm(1766,rep(0,11), rindcor, empirical = TRUE)   create data
> cor(rinddat)
        [,1]   [,2]   [,3]   [,4]   [,5]   [,6]   [,7]   [,8]   [,9]  [,10]  [,11]
 [1,]  1.000 -0.144 -0.244 -0.323 -0.129 -0.056  0.053 -0.043  0.037  0.370  0.186  DADocc
 [2,] -0.144  1.000  0.156  0.088  0.315  0.150 -0.152  0.030  0.035 -0.222 -0.189  RACE
 [3,] -0.244  0.156  1.000  0.274  0.150 -0.039  0.014  0.028  0.002 -0.328 -0.115  NOSIB
 [4,] -0.323  0.088  0.274  1.000  0.218 -0.030 -0.149 -0.060 -0.032 -0.185 -0.118  FARM
 [5,] -0.129  0.315  0.150  0.218  1.000  0.071 -0.292 -0.011 -0.027 -0.211 -0.177  REGN
 [6,] -0.056  0.150 -0.039 -0.030  0.071  1.000 -0.052  0.067  0.018 -0.157  0.111  ADOL
 [7,]  0.053 -0.152  0.014 -0.149 -0.292 -0.052  1.000 -0.010 -0.002 -0.012  0.098  REL
 [8,] -0.043  0.030  0.028 -0.060 -0.011  0.067 -0.010  1.000  0.009 -0.171 -0.122  CIG
 [9,]  0.037  0.035  0.002 -0.032 -0.027  0.018 -0.002  0.009  1.000  0.038  0.216  FECUND
[10,]  0.370 -0.222 -0.328 -0.185 -0.211 -0.157 -0.012 -0.171  0.038  1.000  0.380  ED
[11,]  0.186 -0.189 -0.115 -0.118 -0.177  0.111  0.098 -0.122  0.216  0.380  1.000  AGE
```

var list

```
> #Rindfus model, Freedman page 356

> agereg = tsls(rinddat[,11] ~ rinddat[,10] +  rinddat[,2] +  rinddat[,3] +  rinddat[,4] + [,5]
+    rinddat[,6] +  rinddat[,7] +  rinddat[,8] +  rinddat[,9],  ~  rinddat[,3] + rinddat[,4] [,5,6,7]
+    rinddat[,8] +  rinddat[,9] +  rinddat[,1] +  rinddat[,2])
> summary(agereg)
 2SLS Estimates
Model Formula: rinddat[, 11] ~ rinddat[, 10] + rinddat[, 2] + rinddat[, 3] +
    rinddat[, 4] + rinddat[, 5] + rinddat[, 6] + rinddat[, 7] +
    rinddat[, 8] + rinddat[, 9]
Instruments: ~rinddat[, 3] + rinddat[, 4] + rinddat[, 5] + rinddat[, 6] +
    rinddat[, 7] + rinddat[, 8] + rinddat[, 9] + rinddat[, 1] +
    rinddat[, 2]

                Estimate Std. Error    t value  Pr(>|t|)
(Intercept)   -1.298e-17    0.02083 -6.234e-16 1.000e+00
rinddat[, 10]  4.851e-01    0.08429  5.755e+00 1.020e-08   ED signif
rinddat[, 2]  -1.052e-01    0.02463 -4.270e+00 2.057e-05
rinddat[, 3]   7.692e-02    0.03123  2.463e+00 1.387e-02
rinddat[, 4]  -9.995e-03    0.02380 -4.199e-01 6.746e-01
rinddat[, 5]  -3.628e-02    0.02558 -1.418e+00 1.563e-01
rinddat[, 6]   2.128e-01    0.02415  8.809e+00 0.000e+00
rinddat[, 7]   8.559e-02    0.02324  3.683e+00 2.374e-04
rinddat[, 8]  -5.421e-02    0.02489 -2.178e+00 2.953e-02
rinddat[, 9]   1.966e-01    0.02114  9.300e+00 0.000e+00
```

Ed → Age

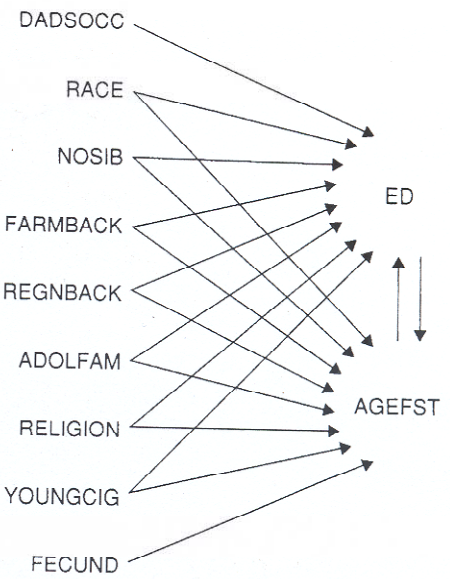
```
> edreg = tsls(rinddat[,10] ~ rinddat[,11] +  rinddat[,2] +  rinddat[,3] +  rinddat[,4] +
+   rinddat[,5] + rinddat[,6] +  rinddat[,7] +  rinddat[,8] +  rinddat[,1],  ~  rinddat[,3]
+   rinddat[,4] +  rinddat[,5] +  rinddat[,6] +  rinddat[,7] +
+   rinddat[,8] +  rinddat[,9] +  rinddat[,1] +  rinddat[,2])
> summary(edreg)
 2SLS Estimates
Model Formula: rinddat[, 10] ~ rinddat[, 11] + rinddat[, 2] + rinddat[, 3] +
    rinddat[, 4] + rinddat[, 5] + rinddat[, 6] + rinddat[, 7] +
    rinddat[, 8] + rinddat[, 1]
Instruments: ~rinddat[, 3] + rinddat[, 4] + rinddat[, 5] + rinddat[, 6] +
    rinddat[, 7] + rinddat[, 8] + rinddat[, 9] + rinddat[, 1] +
    rinddat[, 2]

                Estimate Std. Error    t value  Pr(>|t|)
(Intercept)    4.833e-18    0.01954  2.473e-16 1.000e+00
rinddat[, 11]  1.473e-01    0.09256  1.591e+00 1.118e-01
rinddat[, 2]  -7.652e-02    0.02489 -3.074e+00 2.143e-03
rinddat[, 3]  -2.166e-01    0.02111 -1.026e+01 0.000e+00
rinddat[, 4]  -2.331e-02    0.02182 -1.068e+00 2.856e-01
rinddat[, 5]  -1.093e-01    0.02380 -4.592e+00 4.703e-06
rinddat[, 6]  -1.456e-01    0.02461 -5.917e+00 3.943e-09
rinddat[, 7]  -9.243e-02    0.02103 -4.395e+00 1.173e-05
rinddat[, 8]  -1.278e-01    0.02278 -5.608e+00 2.372e-08
rinddat[, 1]   2.484e-01    0.02469  1.006e+01 0.000e+00

Residual standard error: 0.8213 on 1756 degrees of freedom
```

age not signif

preg ⇏ dropout ?

```
> # ed is significant in age eq, age not signif in ed eq
> #therefore ed --> age  Freedman p.182
```


Figure 1. A Model of the Relationship between Educational Attainment and the Beginning of Motherhood.



$$\widehat{ED} = b_0 + b_1 DADSOCC + b_2 RACE + b_3 NOSIB + b_4 FARMBACK + b_5 REGNBACK + b_6 ADOLFAM + b_7 RELIGION + b_8 YOUNGCIG + b_9 AGEFST + U$$

$$\widehat{AGEFST} = c_0 + c_1 RACE + c_2 NOSIB + c_3 FARMBACK + c_4 REGNBACK + c_5 ADOLFAM + c_6 RELIGION + c_7 YOUNGCIG + c_8 FECUND + c_9 ED + V$$

conomic theory" seems like a natural answer, but an incomplete one. has to be anchored in reality. Sooner or later, invariance needs em- emonstration, which is easier said than done. Outside of economics, tion is perhaps even less satisfactory, because theory is less well de- interventions are harder to define, and the hypothetical experiments kier.

ocial-science example: education and fertility

DAF text

ultaneous equations are often used to model reciprocal causation—U es V, and V influences U. Here is an example. Rindfuss et al (1980) a simultaneous-equations model to explain the process by which a lecides how much education to get, and when to have children. The explanation is as follows.

"The interplay between education and fertility has a significant influ- e on the roles women occupy, when in their life cycle they occupy these s, and the length of time spent in these roles. . . . This paper explores the retical linkages between education and fertility. . . . It is found that the procal relationship between education and age at first birth is dominated he effect from education to age at first birth with only a trivial effect in ther direction.

"No factor has a greater impact on the roles women occupy than mater- Whether a woman becomes a mother, the age at which she does so, and iming and number of subsequent births set the conditions under which r roles are assumed. . . . Education is another prime factor conditioning le roles. . . .

"The overall relationship between education and fertility has its roots at e unspecified point in adolescence, or perhaps even earlier. At this point ations for educational attainment as a goal in itself and for adult roles have implications for educational attainment first emerge. The desire for ation as a measure of status and ability in academic work may encourage en to select occupational goals that require a high level of educational nment. Conversely, particular occupational or role aspirations may set lards of education that must be achieved. The obverse is true for those either low educational or occupational goals. Also, occupational and ational aspirations are affected by a number of prior factors, such as er's education, father's education, family income, intellectual ability, educational experience, race, and number of siblings. . . ."

model used by Rindfuss et al (the paper is reprinted at the back of consists of two linear equations in two unknowns, ED and AGE:

$$\text{ED} = a\text{AGE} + A_i + \delta_i,$$
$$\text{AGE} = a'\text{ED} + A_i' + \epsilon_i.$$

 Analysis on handout "C"

According to the model, a woman—indexed by the subscript i—chooses her educational level ED_i and age at first birth AGE_i as if by solving the two equations for the two unknowns.

These equations are response schedules (section 4.5). The coefficients a and a' are parameters, to be estimated from the data. (Here, the prime doesn't denote a transpose: a is scalar.) The terms A_i and A_i' take background factors into account:

$$A_i = a_0 + b\text{OCC}_i + c_1\text{RACE}_i + \cdots + c_7\text{YCIG}_i, \tag{19}$$
$$A_i' = a_0' + b'\text{FEC}_i + c_1'\text{RACE}_i + \cdots + c_7'\text{YCIG}_i. \tag{20}$$

Variables are defined in table 1, and the notes to the table describe the sample survey that collected the data.

The parameters $a_0, b, c_1, \ldots$ are to be estimated from the data. The random errors (δ_i, ϵ_i) are assumed to have mean 0, and (as pairs) to be indepen-

Table 1. Variables in the model (Rindfuss et al 1980).

The endogenous variables	
ED	Respondent's education (Years of schooling completed at first marriage)
AGE	Respondent's age at first birth
The exogenous variables	
OCC	Respondent's father's occupation
RACE	Race of respondent (Black = 1, other = 0)
NOSIB	Respondent's number of siblings
FARM	Farm background (coded 1 if respondent grew up on a farm, else coded 0)
REGN	Region where respondent grew up (South = 1, other = 0)
ADOLF	Broken family (coded 0 if both parents present when respondent was 14, else coded 1)
REL	Religion (Catholic = 1, other = 0)
YCIG	Smoking (coded 1 if respondent smoked before age 16, else coded 0)
FEC	Fecundability (coded 1 if respondent had a miscarriage before first birth; else coded 0)

Notes: The data are from a probability sample of 1766 women 35–44 years of age residing in the continental United States. The sample was restricted to ever-married women with at least one child. OCC was measured on Duncan's scale (section 5.1), combining information on education and income. Notation differs from Rindfuss et al.

dent and identically distributed from woman to woman. The model allows δ_i and ϵ_i to be correlated; δ_i may have a different distribution from ϵ_i.

Rindfuss et al use two-stage least squares to fit the equations. Notice that they have excluded FEC from equation (19), and OCC from equation (20). Without these identifying restrictions, the system would be under-identified (section 2 above).

The main empirical finding is this. The estimated coefficient of AGE in (17) is not statistically significant, i.e., a could be zero. The woman who dropped out of school because she got pregnant at age 16 would have dropped out anyway: the causal arrow points from ED to AGE, not the other way.

This finding depends on the model. When looked at coldly, the argument may seem implausible. A critique can be given along the following lines.

(i) Assumptions about the errors. Why are the errors independent and identically distributed across the women? Independence may be reasonable, but heterogeneity is more plausible than homogeneity.

(ii) Omitted variables. Important variables have been omitted from the model, including two that were identified by Rindfuss et al themselves—aspirations and intellectual ability. (See the quotes at the beginning of the section.) Since Malthus (1798), it has been considered that wealth is an important factor in determining education and marriage. Wealth is not in the model. Social class matters, and OCC measures only one of its aspects.

(iii) Why additive linear effects?

(iv) Constant coefficients. Rindfuss et al are assuming that the same parameters apply to all women alike, from poor blacks in the cities of the Northeast to rich whites in the suburbs of the West. Why?

(v) Are FEC and OCC exogenous?

(vi) What about the identifying restrictions?

(vii) Are the equations structural?

It is easier to think about questions (v–vii) in the context of a model that restricts attention to a more homogeneous group of women, where the only relevant background factors are OCC and FEC. The equations—response schedules—are as follows.

$$\text{ED} = a_0 + a\text{AGE} + b\text{OCC} + \delta_i, \tag{21}$$

$$\text{AGE} = a_0' + a'\text{ED} + b'\text{FEC} + \epsilon_i\,. \tag{22}$$

What do these equations tell us? Two hypothetical experiments help answer this question. In both experiments, fathers are assigned to jobs; and

daughters are assigned to have a miscarriage before giving birth to their first child (FEC = 1), or not to have a miscarriage (FEC = 0).

> Experiment #1. Daughters are assigned to the various levels of AGE. ED is observed as the response. In other words, the hypothetical experimenter chooses when the woman has her first child, but allows her to decide when to leave school.
>
> Experiment #2. Daughters are assigned to the various levels of ED. Then AGE is observed as the response. The hypothetical experimenter decides when the woman has had enough education, but lets her have a baby when she wants to.

The statistical terminology is rather dry. The experimenter makes fathers do one job rather than another: surgeons cut pastrami sandwiches and taxi drivers run the central banks. Women are made to miscarry at one time and have their first child at another.

The equations can now be translated. According to (21), in the first experiment, ED does not depend on FEC. (That is one of the identifying restrictions assumed by Rindfuss et al.) Moreover, ED depends linearly on AGE and OCC, plus an additive random error. According to (22), in the second experiment, AGE does not depend on OCC. (That is the other identifying restriction assumed by Rindfuss et al.) Moreover, AGE depends linearly on ED and FEC, plus an additive random error. Even for thought experiments, this is a little fanciful.

We return now to the full model, equations (17–20). The data were collected in a sample survey, not an experiment (notes to table 1). Rindfuss et al must be assuming that Nature assigned OCC, RACE, . . . , FEC independently of the disturbance terms δ and ϵ in (17) and (18). That assumption is what makes OCC, RACE, . . . , FEC exogenous. Rindfuss et al must further be assuming that women chose ED and AGE as if by solving the two equations (17) and (18) for the two unknowns, ED and AGE. Without this assumption, simultaneous-equation modeling seems irrelevant. (The comparable element in the butter model is the law of supply and demand.)

The equations estimated from the survey data must apply as well to experimental situations where ED and AGE are manipulated. For instance, women who freely choose their educational levels and times to have children do so using the same pair of equations—with the same parameter values and error terms—as women made to give birth at certain ages. The data analysis in the paper doesn't justify such assumptions: how could it? But these constancy assumptions are the basis for causal inference from non-experimental data.

Without the response schedules, it is hard to see what "effects" might mean, apart from slopes of some plane that has been fitted to survey data. It would remain unclear why the plane should be fitted by two-stage least squares, or what role the significance tests are playing. Rindfuss et al have an interesting question, and there is much wisdom in their paper. But they have not demonstrated a connection between the social problem they are studying and the statistical technique they are using.

Simultaneous equations that derive from response schedules are structural. Structural equations describe real observational studies, and the hypothetical experiments that usually remain behind the scenes. Unless equations are structural, they have no causal implications (section 5.5).

More on Rindfuss et al

Rindfuss et al have arguments to support their position, but these are not convincing. For instance, exogeneity is discussed in the paper, and in Rindfuss and St. John (1983). However, the discussion misses the critical point: variables labelled as "instrumental" or "exogenous," like OCC, RACE, ..., FEC, need to be independent of the error terms. Why would that be so? Moreover, justifications given for the identifying restrictions seem artificial.

Hofferth and Moore (1979, 1980) obtain different results using different instruments, as noted by Hofferth (1984). Rindfuss et al (1984) say that

> "instrumental variables.... require strong theoretical assumptions.... and can give quite different results when alternative assumptions are made.... it is usually difficult to argue that behavioral variables are truly exogenous and that they affect only one of the endogenous variables but not the other." [pp. 981–82]

Thus, results depend quite strongly on assumptions about identifying restrictions and exogeneity, and there is no good way to justify one set of assumptions rather than another. Also see Bartels (1991), who comments on the impact of exogeneity assumptions and the difficulty of verification. Rindfuss and St. John (1983) give useful detail on the model. There is an interesting exchange between Geronimus and Korenman (1993) and Hoffman et al (1993) on the costs of teenage pregnancy.

8.6 Covariates

In the butter hypothetical, we could take the exogenous variables as non-manipulable covariates. The assumption would be that Nature chooses

$$(W_t, H_t, T_t, M_t) : t = 1, \ldots, 20$$

independently of the random error terms

$$(\delta_t, \epsilon_t) : t = 1, \ldots, 20.$$

The error terms would still be assumed IID (as pairs) with mean 0, and a 2×2 covariance matrix. We still have two hypothetical experiments: (i) set the price P to farmers, and see how much butter comes to market; (ii) set the price P to consumers and see how much butter is bought. By assumption, the answer to (i) is still

$$Q = a_0 + a_1 P + a_2 W_t + a_3 H_t + \delta_t, \tag{23}$$

while the answer to (ii) is

$$Q = b_0 + b_1 P + b_2 T_t + b_3 M_t + \epsilon_t. \tag{24}$$

For the observational data, we would still need to assume that Q_t and P_t in year t are determined as if by solving (23) and (24) for the two unknowns, Q and P, which gets us back to (2a) and (2b).

With Rindfuss et al, OCC, RACE, ..., FEC could be taken as non-manipulable covariates, eliminating some of the difficulty in the hypothetical experiments. The identifying restrictions—FEC is excluded from (19) and OCC from (20)—remain mysterious, as does the assumed linearity. How could you verify such assumptions?

Terminology. Often, "covariate" just means a right hand side variable in a regression equation—especially if that variable is only included to control for a possible confounder. Sometimes, "covariate" signifies a non-manipulable characteristic, like age or sex. Non-manipulable variables are occasionally called "concomitants."

8.7 Linear probability models

Schneider et al (1997) use two-stage least squares—with lots of bells and whistles—to study the effects of school choice on social capital. (The paper is reprinted at the back of the book.) "Linear probability models" are used to control for confounders and self-selection. The estimation strategy is quite intricate. Let's set the details aside, and think about the logic. First, here is what Schneider et al say they're doing, and what they found:

> "While the possible decline in the level of social capital in the United States has received considerable attention by scholars such as Putnam and Fukuyama, less attention has been paid to the local activities of citizens that help define a nation's stock of social capital .... giving parents greater choice over the public schools their children attend creates incentives for parents as

Weeks 6 [redacted]

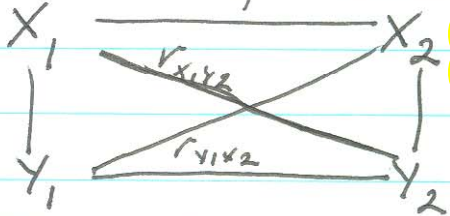
Stat 209 [redacted]

Reciprocal Effects, Simultaneous Equations

CLC (cross-lagged correlation)

compare $r_{X_1 Y_2}$ to $r_{Y_1 X_2}$

causal predominance to the larger (sig)



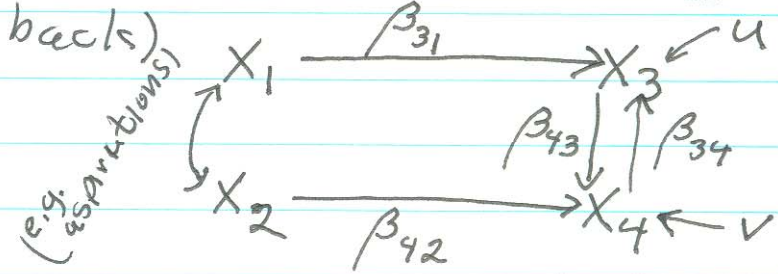
Longitudinal data

Non-Recursive (feedback)

$X_3 = \beta_{31} X_1 + \beta_{34} X_4 + u$

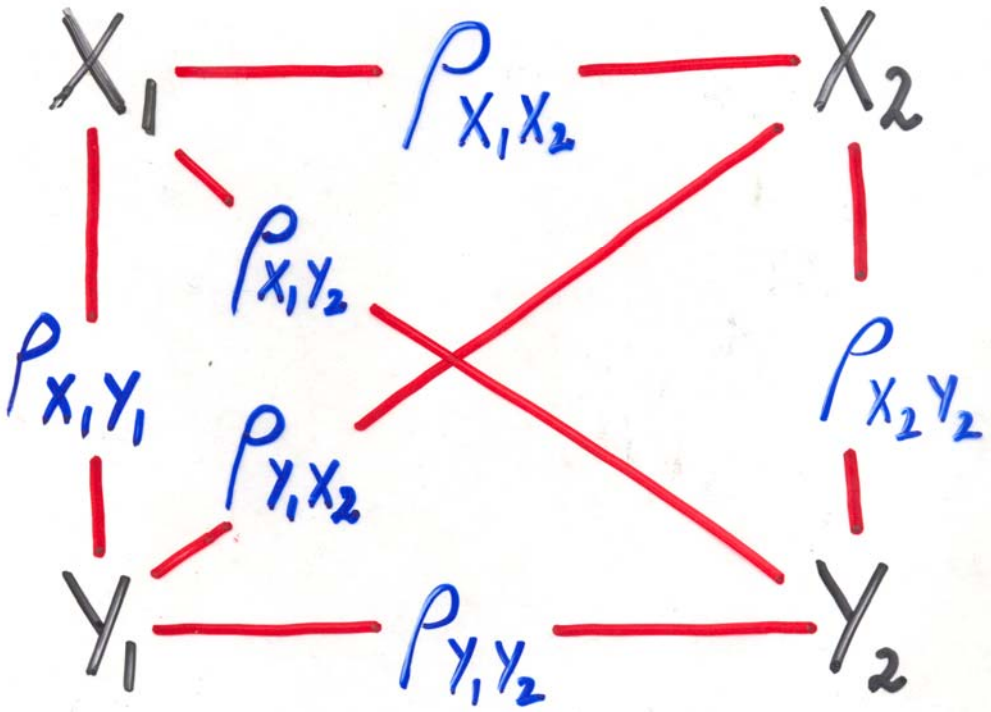
$X_4 = \beta_{42} X_2 + \beta_{43} X_3 + v$

$E(X_1) = 0 \quad E(u) = E(v) = 0$

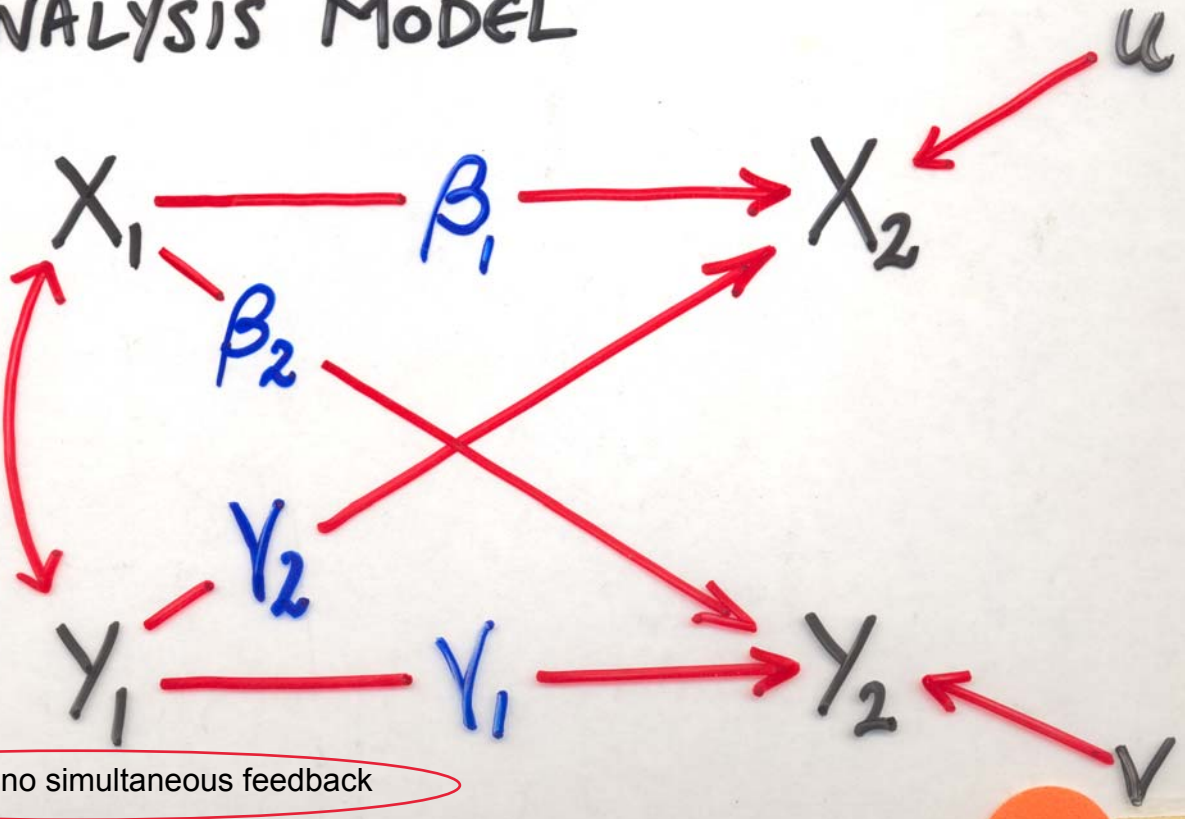


RECIPROCAL EFFECTS

CROSS-LAGGED CORRELATION "MODEL"



PATH ANALYSIS "MODEL"



effects are lagged, no simultaneous feedback

Cross-lagged credo

"With the introduction of the cross-lagged panel correlation method ..., causal inferences based on correlational data obtained in longitudinal studies can be made and enjoy the same logical status as those derived in the more standard experimental settings."

Crano and Mellon (1978) p. 41

Reciprocal Effects $X \rightleftarrows Y$

Longitudinal TV violence, aggression

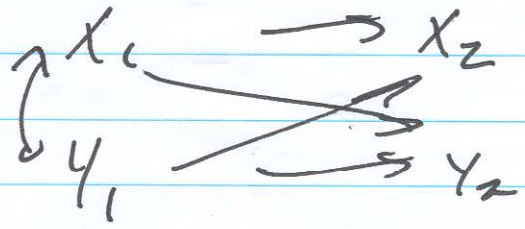
CLC $r_{X_1 Y_2}$ vs $r_{Y_1 X_2}$

Linked differential eq's

$$\frac{dY}{dt} = a + bX + cY$$

$$\frac{dX}{dt} = d + eX + fY$$

$\Rightarrow$ lagged regression discrete time



Time-series, Granger causality Pierce + Haugh.

discrete data Lazarsfeld 16-fold table (Lazarsfeld + Merton)

Simultaneous (reciprocal) effects

no lagged repeated measures

USE IV exogenous vars to estimate relative effects

Duncan Haller Portes peer influences

Rindfuss (Jencks) ed, fertility (cohort)

the context of particular hypothesized models, and the assumptions contained in these models must be defended in a given situation. In particular, two-wave reciprocal effects models involve a set of possibly restrictive assumptions that can be relaxed only, as in cross-sectional research, by including outside variables. Three-wave and multiwave panels, however, can be estimated by imposing fewer constraints on the causal parameters.

Cross-Lagged Effects Models

The Two-Wave Model

The most basic model for estimating possible reciprocal effects is an extension to two dependent variables of the lagged effect static-score model considered in the previous chapter (Equation 2.14), with each variable at time 2 being predicted by its previous value as well as the time 1 value of the other variable of interest. This model is shown in diagram form in Figure 3.1. X_1 and X_2 represent one variable (e.g., group memberships) measured across the two panel waves, and Y_1 and Y_2 another variable (e.g., protest potential) measured at both time points. X_2 and Y_2 are hypothesized to be determined by their wave 1 values, the lagged value of the other variable, and an error term U. The correlation between the wave 1 variables is represented in the figure by ρ_1, and the correlation between the structural disturbances of the wave 2 equations is represented by $\rho_{U_1U_2}$. The two structural equations can be written as follows:

$$\begin{aligned} Y_2 &= \beta_1 X_1 + \beta_2 Y_1 + U_1 \\ X_2 &= \beta_3 Y_1 + \beta_4 X_1 + U_2 \end{aligned} \tag{3.1}$$

with all variables in the model expressed in mean deviation form to eliminate consideration of the intercept term.

The cross-lagged model has wide applicability in panel analysis. When change in the dependent variables is modeled according to a discrete time process, the cross-lagged model will be appropriate whenever the causal lags are approximately equal to the time period between measurements. When change in the dependent variables occur continuously, it can be shown that the cross-lagged model represents the integral solution to a system of differential equations where the instantaneous rates of change in X and Y are dependent on one another over time, as in

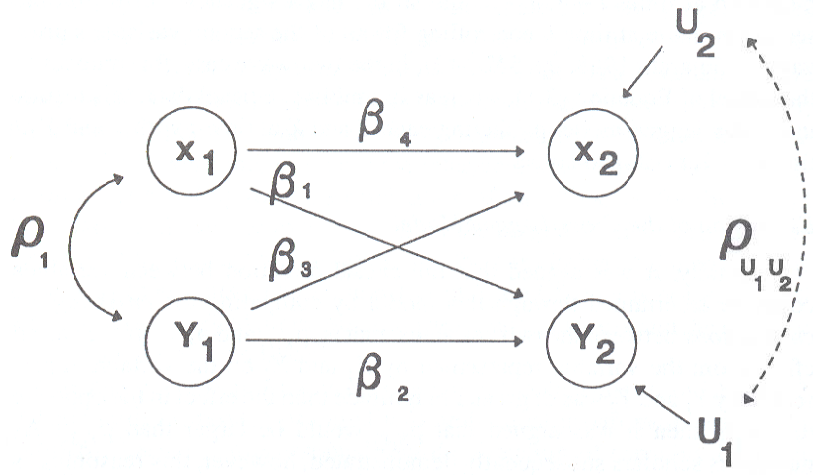


Figure 3.1. Two-Wave Model With Cross-Lagged Effects

$$\begin{aligned} \frac{dY_t}{dt} &= c_0 + c_1 X_t + c_2 Y_t \\ \frac{dX_t}{dt} &= c_3 + c_4 Y_t + c_5 X_t . \end{aligned} \tag{3.2}$$

The β_k in Equation (3.1) are, as in the single-equation continuous time model considered in the previous chapter, nonlinear functions of the c coefficients, the "fundamental parameters of change" in Equation (3.2), and the time between panel waves. The calculation of the c coefficients from the β_k involves complex mathematical manipulations beyond the scope of this monograph; interested readers are referred to Arminger (1986), Coleman (1968, pp. 448-452), and Tuma and Hannan (1984, Chapters 11-12) for details. For our purposes, it is sufficient to note that when the causal system is one of continuous reciprocal feedback from one variable to another, the cross-lagged model "tends not to be misleading about the *direction* of causal influence" (Dwyer, 1983, p. 352).

It may also be noted that the cross-lagged model corresponds to the so-called "Granger test" for causality in time series analysis, which posits that a variable "Granger-causes" the other if any value of the first variable

DOES TELEVISION VIOLENCE CAUSE AGGRESSION? [1]

LEONARD D. ERON [2]
University of Illinois at Chicago Circle
MONROE M. LEFKOWITZ
New York State Department of Mental Hygiene, Albany
L. ROWELL HUESMANN
Yale University
LEOPOLD O. WALDER
Organization for Research in the Behavioral Sciences, Greenbelt, Maryland

WITH the increasing prominence of violence in our society, social scientists have been turning their attention to the antecedents of aggressive behavior in children and adults. Television programming with its heavy emphasis on interpersonal violence and acquisitive lawlessness has been assigned a role both in inciting aggression and teaching viewers specific techniques of aggressive behavior. The relation between overt aggression and television habits has been demonstrated in a few survey studies which, however, because of the nature of surveys have not been able to discriminate cause and effect (Bailyn, 1959; Eron, 1963; Schramm, Lyle, & Parker, 1961).

On the other hand, manipulative laboratory experiments have demonstrated an immediate effect on the extent of aggressive behavior of subjects who have witnessed aggressive displays on film (Bandura, Ross, & Ross, 1963a; Berkowitz, Corwin, & Heironimus, 1963). The latter studies, however, can be criticized for not duplicating real-life television viewing situations and possibly not accounting for anything more than a transient effect on the viewer. Hartup and Yonas (1971) stated in their review of developmental psychology in the latest *Annual Review of Psychology,*

> Current studies of the childhood determinants of aggression are not extensive. The report of a presidential commission, *Violence and the Media* (Baker & Ball, 1969), indicates that child psychologists possess much information concerning the determinants of aggression but very little stemming directly from naturalistic sources [p. 375].

One possible way of utilizing survey procedures to demonstrate cause and effect is to use a longitudinal context. By contrasting the magnitude of contemporaneous and longitudinal correlations between two sets of variables, it is possible to account more clearly for which of the variables is antecedent and which consequent. The authors have now accumulated data on both aggressive behavior and television viewing habits over a 10-year period in a large group of subjects first seen when they were 8–9 years of age. Thus, we can implement such an analysis.

The hypotheses of this research are that a young adult's aggressiveness is positively related to his preference for violent television when he was 8–9 years old and, furthermore, that his preference for violent television during this critical period is one cause of his aggressiveness.

METHOD

Longitudinal data were collected on 427 teenagers of an original group of 875 children who had participated in a study of third-grade children in 1960 (Eron, 1963; Eron, Walder, & Lefkowitz, 1971). The original 875 constituted the entire third-grade population of a semirural county in New York's Hudson River Valley, while the 427 subjects were those who could be located and interviewed 10 years later.

The information collected about these subjects in both time periods falls into two classes: (*a*) measures of aggression and (*b*) potential predictors of aggression. During the third-grade interviews, four different data sources had been used: the subject, his peers, his mother, and his father. Ten years later, the data sources were the subject and his peers. For convenience, this later time period will be designated as the thirteenth grade.

The variables used in the study are listed in Table 1. Two variables are of particular importance: peer-rated

[1] This research was supported by Grant M1726 from the National Institute of Mental Health and Contract No. HSM 42-70-60 from the Surgeon General's Scientific Advisory Committee on Television and Social Behavior. Thanks are also due to the New York State Department of Mental Hygeine for their support and to Anne Karabin, Marjorie Kline, Ann McAleer, Victor Pompa, and Ann Yeager who served as research assistants.

[2] Requests for reprints should be sent to Leonard D. Eron, Department of Psychology, University of Illinois, Box 4348, Chicago, Illinois 60680.

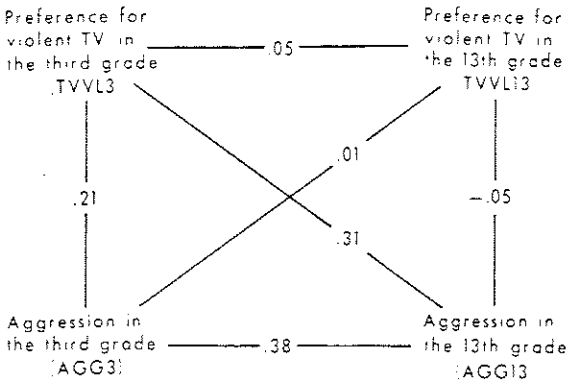


FIG. 1. The correlations between a preference for violent television and peer-rated aggression for 211 boys over a 10-year lag.

television in the third grade and aggression in the thirteenth grade, whether the measure is one of peer-rated aggression or self-ratings of aggression (antisocial behavior and MMPI–49S).

Causal Analysis

Having established that there exists a highly significant relationship between a preference for violent television in the third grade and aggressive habits in the thirteenth grade, one can consider the alternative causal explanations for this phenomenon. Of course, one cannot demonstrate that a particular hypothesis is true. One can only reject untenable hypotheses and present evidence on the plausibility of the remaining hypotheses.

Cross-lagged correlations. Consider the pattern of correlations diagramed in Figure 1; the correlations on the diagonals are called cross-lagged correlations. The cross-lagged correlation between a preference for violent television in the third grade and thirteenth-grade aggression was highly significant. When coupled with the lack of a relation between third-grade aggression and a preference for violent television in the thirteenth grade, this significant correlation supports the hypothesis that preferring to watch violent television is a cause of aggressive behavior. This causal hypothesis is diagramed in Figure 2a. The probability of a chance occurrence of the difference between the cross-lagged correlations is low (Fisher's $z = 3.07$, $p < .002$); however, a few rival hypotheses are seemingly consistent with the difference and deserve consideration.

One alternative hypothesis is that preference for violent television in the third grade stimulates concurrent aggression, and this aggression leads to thirteenth-grade aggression or at least is being remeasured in the thirteenth grade. The corresponding causal chain is diagramed in Figure 2b. This interpretation can be rejected because if it were true, the relation between the end points of the causal sequence would have been no stronger than the product of the relations between all adjacent intermediate points. But the correlation between the end points was .31, which was much higher than the product of the intermediate correlations.

For a similar reason, the causal chain diagramed in Figure 2c can be eliminated as an alternative hypothesis. If early aggression caused a preference for violent television which in turn contributed to later aggression, the correlation between early and later aggression would have been less than the product of the two intermediate correlations. It was not.

One cannot reject so easily the more realistic alternative hypothesis diagramed in Figure 2d. This causal hypothesis asserts that early aggression causes both contemporaneous preferences for violent television and later aggression. Part of this hypothesis, that early aggression contributes to later aggression, is quite probably true. What is of interest here, though, is whether or not the relation between early television preferences and later aggression can be explained as an artifact of early aggression. One can obtain evidence to refute this idea by

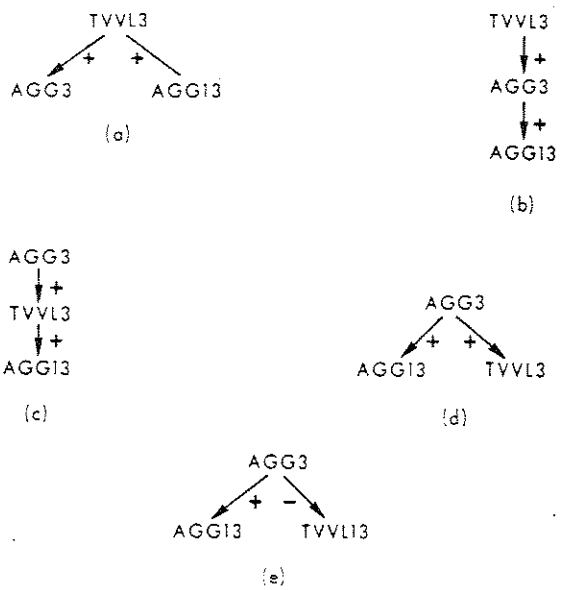


FIG. 2. Five feasible causal hypotheses for the correlations presented in Figure 1.

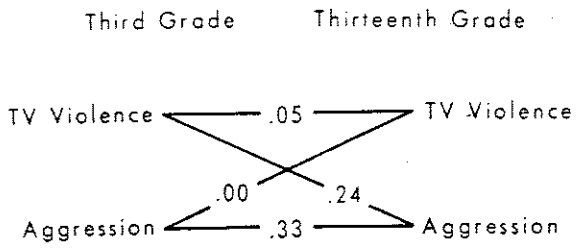


FIG. 3. The path coefficients for a preference for watching violent television and peer-rated aggression.

life independent of the other causal contributors studied.

Path analysis. A more specialized technique for using multiple regression coefficients to estimate causal effects is path analysis (Heise, 1970). The path coefficients for television and aggression are shown in Figure 3. These coefficients are standardized partial regression coefficients. In other words, the path coefficient from third-grade television violence to thirteenth-grade aggression is the coefficient of third-grade television violence in a regression equation predicting thirteenth-grade aggression with third-grade aggression controlled. The obtained pattern of path coefficients adds further credence to the argument that watching violent television contributes to the development of aggressive habits.[3]

DISCUSSION

The above results indicate that television habits established by age 8–9 years influence boys' aggressive behavior at that time and at least through late adolescence. The more violent are the programs preferred by boys in the third grade, the more aggressive is their behavior both at that time and 10 years later. This relation between early television habits and later aggression prevails both for peer-rated aggression and for self-ratings of aggression. Actually, these early television habits seem to be more influential than current viewing patterns since a preference for violent television in the thirteenth grade is not at all related to concurrent aggressive behavior, nor are early television habits related to later television habits.

It would be very difficult to explain these results as methodological artifacts. While the peer-rated aggression measure possesses demonstrated reliability (Walder et al., 1961), the thirteenth-grade scores represent more than the temporal reliability of the third-grade measure. Not only had 10 years of behavior intervened between the measurement periods, but the set of raters of each subject was substantially different as well. It is also unlikely that the findings were produced by a common method rather than content. Method refers to the form of the measuring device and the procedure for obtaining and calculating the score; content refers to what one intends to measure. Feshbach (1970), for example, warned that the

> predictive utility of sociometric, projective, and inventory measures is limited by the substantial method variance yielded by each procedure, and it is evident that the more dissimilar the test of aggression is to the aggression criterion, the weaker are the relationships obtained [p. 181].

The peer-rated aggression measure employed in this study should not be susceptible to such criticism as it has been extensively validated over a 10-year period using a variety of techniques ranging from Campbell and Fiske's (1959) multitrait, multimethod technique to factor analysis. Peer-rated aggression scores agree closely with teacher ratings, clinical referrals, and overt behavior measured by such devices as the Iowa Aggression Machine (Williams, Meyerson, Eron, & Semler, 1967).

The ratings of the television programs for violence content also possess demonstrated reliability, as noted above, and were in close agreement with ratings made by at least two other groups (Feshbach & Singer, 1971; Greenberg & Gordon, 1970).

While only preferences for violent television were measured in both third and thirteenth grades, it is reasonable to assume that a child's preference for a television program is very highly correlated with the length of time that he attends to that program. The report of television preference in the third grade was made by the mother who probably was strongly influenced by her child's actual viewing habits, especially since the question regarding television preference was asked immediately following the question concerning the length of time the child watched television. Hence, it seems fair to conclude that a preference for violent television is indicative of viewing violent television by the child, particularly in the third grade. Furthermore, since the programs mentioned were the preferred ones, not only is it likely that they watched these

[3] The authors are indebted to John M. Neale, State University of New York at Stony Brook, for this path analysis of our data and to David Kenny, Northwestern University, for suggestions regarding interpretation of the cross-lagged correlations.

Psychological Bulletin
1980, Vol. 88, No. 2

A Critique of Cross-Lagged Correlation

David Rogosa
Department of Education, University of Chicago

Cross-lagged correlation is not a useful procedure for the analysis of longitudinal panel data. In particular, the difference between the cross-lagged correlations is not a sound basis for causal inference. Demonstrations of the failure of cross-lagged correlation are based mainly on results for the two-wave, two-variable longitudinal panel design. Extensions of these results to panels with multiple waves and multiple measures reveal additional problems.

The topic of this article is the analysis of reciprocal causal effects. Often, questions about reciprocal causal effects have been phrased, Does X cause Y or does Y cause X? More formally, many have spoken of a determination of causal predominance or of a preponderant causal effect. Examples of such research questions in developmental psychology include the reciprocal influences in mother–child interaction (Clarke-Stewart, 1973) and relationships between infant intelligence and infant behavior (Crano, 1977). Examples from educational research include the relationship between teacher expectation and student achievement (Crano & Mellon, 1978; Humphreys & Stubbs, 1977; West & Anderson, 1976) and the relationship between self-concept and achievement (Bachman & O'Malley, 1977; Calsyn & Kenny, 1977; Purkey, 1970). Empirical research on topics such as these has resulted in the collection and analysis of large amounts of longitudinal panel data.

Data from a longitudinal panel consist of observations on n cases at T ($t = 1, \ldots, T$) time points or waves. At each time point, observations on one or more variables are obtained. Much attention is given to the simplest relevant panel design: the two-wave, two-variable (2W2V) longitudinal panel. For 2W2V panels the two variables are labeled X and Y. These variables are subscripted to indicate the time of measurement. Thus for each individual case, measures X_1, Y_1, X_2, and Y_2 are available.

Cross-lagged correlation (CLC) is currently the most popular procedure in many areas of psychological and educational research for identifying causal effects from longitudinal panel data. Most often CLC is used to determine a predominant causal influence—the causal winner.

Users of CLC often make enthusiastic claims. For example, Crano and Mellon (1978) asserted,

> With the introduction of the cross-lagged panel correlational method. . ., causal inferences based on correla-

This research was supported by a seed grant from the Spencer Foundation. The assistance of David Brandt and Michele Zimowski in this research is gratefully acknowledged.

Requests for reprints should be sent to David Rogosa, who is now at the School of Education, Stanford University, Stanford, California 94305.

 0033-2909/80/8802-0245$00.75

be an improvement over current practice in CLC; the construction of the interval estimate can be found in Olkin (1966).

Also, significant differences between the cross-lagged correlations are interpreted without regard for the assumptions of CLC. Rarely is any assessment of stationarity made in applications of CLC. In many sets of longitudinal data, the difference between the synchronous correlations is about as large as the difference between the cross-lagged correlations (e.g., Humphreys & Stubbs, 1977, Table 4; Kenny, 1975, Tables 3 and 5). In a few applications a test of the null hypothesis—$H_0: \rho_{X_1Y_1} = \rho_{X_2Y_2}$—is performed. However, nonrejection does not prove the null hypothesis true, and in addition this null hypothesis is only a necessary condition for stationarity. The statistical test of equal cross-lagged correlations ignores the assumptions and the use of a preliminary test for stationarity. The complete null hypothesis should be that the cross-lagged correlations are equal, conditional on stationarity. That is, the conditional null hypothesis of interest is $H_0: \rho_{X_1Y_2} = \rho_{Y_1X_2} | \rho_{X_1Y_1} = \rho_{X_2Y_2}$. No exact statistical test is available for this conditional null hypothesis, although methods such as covariance structure analysis (Jöreskog & Sörbom, 1979) could be used to form a large-sample, normal-theory test. Of course, no improvement in the use of statistical inference procedures can offset the basic deficiencies of CLC.

Summary and Discussion

No justification was found for the use of CLC. In CLC both determinations of spuriousness and attributions of causal predominance are unsound. The results for 2W2V panels demonstrate that when reciprocal causal effects are absent, the difference between the cross-lagged correlations may be either small or large, and when reciprocal causal effects are present, the difference between the cross-lagged correlations may be either small or large. Also, the practice in CLC of reducing the analysis of data with multiple waves and multiple measures to a collection of 2W2V analyses produces additional problems. CLC is best forgotten.

It should be stressed that this article is not devoted to identifying perverse situations in which CLC might break down. Rather, a straightforward and explicit formulation of causal effects in panel data showed that CLC is not a sound basis for causal inference over a wide range of plausible situations. Possibly, CLC could be patched up, primarily through additional restrictive assumptions, in response to the deficiencies demonstrated in this article. This article is not intended to stimulate such activities. CLC should be set aside as a dead end.

In some sense this article is a flight of fantasy. The fantasy is the notion of a closed two-variable causal system on which the exposition and results are based. This simple formulation serves well for investigating the worth of CLC. Also, almost all applications and technical development of CLC have been limited to two-variable causal systems. CLC fails even in this idealized situation, and no grounds for optimism exist for better performance in more complex causal systems.

In his articles on longitudinal panel data, Duncan (1969, 1972, 1975) stressed that the analysis of panel data cannot be reduced to a mechanical procedure that yields trustworthy inferences about causal structures. A minimal requirement for success is the careful formulation of explicit (and often specialized) models for the substantive processes. An intent of this article is to emphasize this message. Trying to answer a causal question from a set of (longitudinal) data is asking a lot from those data. Minimal requirements are that the right variables be measured well. Often the state of theoretical and empirical knowledge in a substantive area is not sufficiently advanced that the relevant variables have been identified or that sufficient measurement techniques have been developed.

Alternative methods for analyzing panel data were not endorsed or discussed in detail. This omission reflects the fact that methods for detecting patterns of causal influence from panel data are far from fully developed. The contribution of this article is to demonstrate that CLC certainly is not the method to rely on for the analysis of panel data. It is hoped that this article will direct efforts away from further development and application of CLC and toward the development and evaluation of productive approaches for the analysis of panel data.

It seems appropriate to make the point that social scientists frequently have been attracted to methods for the analysis of nonexperimental data that are far more flawed and less justified than path analysis and relatives. A most vivid example is provided by the method of cross-lagged correlation, which remains a very popular procedure for the analysis of reciprocal effects from nonexperimental, longitudinal data. Cross-lagged correlation purports to answer the question—Does X cause Y or does Y cause X?—by a simple comparison of the lagged correlations between X and Y (i.e., the correlations between X_1 and Y_2 and Y_1 and X_2 for two time points). A remarkable attribution of as-if-by-experiment is provided by Crano and Mellon (1978): "With the introduction of the cross-lagged panel correlation method . . . , causal inferences based on correlational data obtained in longitudinal studies can be made and enjoy the same logical status as those

From: Casual Models.....

derived in the more standard experimental settings" (p. 41). In other words, the use of cross-lagged correlation dispenses with the need for experiments, statistical models, or careful data analysis; a quick comparison of a few correlation coefficients is all that's required to study reciprocal effects.

Rogosa (1980) was only one in a tradition of papers, starting with Duncan (1969), Goldberger (1971), and Heise (1970), sharply critical of cross-lagged correlation. Even Cook and Campbell (1979, Chap. 7) are unenthusiastic about the usefulness of cross-lagged correlation; yet most advocates and users of this procedure remain undaunted. Rogosa (1980) exposits a number of simple statistical models for reciprocal effects between two variables—path analysis models, continuous-time feedback models, and multiple time series models. The mathematical results in Rogosa (1980) demonstrate the inability of the method of cross-lagged correlation to recover the structure of the reciprocal effects specified by these models. Results and numerical examples are presented for two-wave and multiwave data. Rogosa (1985) provides a nontechnical overview and extensive references on approaches to the analysis of reciprocal effects.

The mathematical and numerical demonstrations of the failures of cross-lagged correlation in Rogosa (1980) had the following simple, limited structure. Start with a basic path-analysis regression model for two variables X and Y measured at times 1 and 2 (the popular two-wave, two-variable panel design)

$$X_2 = \beta_0 + \beta_1 X_1 + \gamma_2 Y_1 + u,$$
$$Y_2 = \gamma_0 + \beta_2 X_1 + \gamma_1 Y_1 + v. \qquad (4)$$

In the context of the statistical model in (4) the parameters β_1 and γ_1 represent the influence of a variable on itself over time. The parameters β_2 and γ_2 represent the lagged, reciprocal effects between X and Y; thus the relative magnitudes of β_2 and γ_2 are presumed to indicate the nature of the reciprocal causal effects. In Rogosa (1980) combinations of β_2 and γ_2 values are compared with the results of the method of cross-lagged correlation. The major (and perhaps only) virtue of the path analysis model (4) is the identification of specific parameters believed to represent the reciprocal effects. If this model of the reciprocal influences between X and Y were valid, then estimation of β_2 and γ_2 would inform about reciprocal effects. Perhaps the best way to think about (4) and the related structural regression models is that these comprise a simple statistical model for reciprocal effects that, however, may be a far-from-satisfactory scientific model of the psychological (etc.) process.

The real moral about the analysis of reciprocal effects is that you can't estimate something without first defining it, and statistical models at least allow definition of key parameters. Regrettably, the seductive simplicity of cross-lagged correlation has inhibited serious work on the complex question of reciprocal effects. Despite the complexity of research questions about

reciprocal effects, empirical research has attempted to answer the over-simplified question, Does X cause Y or does Y cause X? by casually comparing a couple of correlations.

JAMA Pediatrics | Original Investigation

Association Between Screen Time and Children's Performance on a Developmental Screening Test

Sheri Madigan, PhD; Dillon Browne, PhD; Nicole Racine, PhD; Camille Mori, BA; Suzanne Tough, PhD

IMPORTANCE Excessive screen time is associated with delays in development; however, it is unclear if greater screen time predicts lower performance scores on developmental screening tests or if children with poor developmental performance receive added screen time as a way to modulate challenging behavior.

OBJECTIVE To assess the directional association between screen time and child development in a population of mothers and children.

DESIGN, SETTING, AND PARTICIPANTS This longitudinal cohort study used a 3-wave, cross-lagged panel model in 2441 mothers and children in Calgary, Alberta, Canada, drawn from the All Our Families study. Data were available when children were aged 24, 36, and 60 months. Data were collected between October 20, 2011, and October 6, 2016. Statistical analyses were conducted from July 31 to November 15, 2018.

EXPOSURES Media.

MAIN OUTCOMES AND MEASURES At age 24, 36, and 60 months, children's screen-time behavior (total hours per week) and developmental outcomes (Ages and Stages Questionnaire, Third Edition) were assessed via maternal report.

RESULTS Of the 2441 children included in the analysis, 1169 (47.9%) were boys. A random-intercepts, cross-lagged panel model revealed that higher levels of screen time at 24 and 36 months were significantly associated with poorer performance on developmental screening tests at 36 months (β, −0.08; 95% CI, −0.13 to −0.02) and 60 months (β, −0.06; 95% CI, −0.13 to −0.02), respectively. These within-person (time-varying) associations statistically controlled for between-person (stable) differences.

CONCLUSIONS AND RELEVANCE The results of this study support the directional association between screen time and child development. Recommendations include encouraging family media plans, as well as managing screen time, to offset the potential consequences of excess use.

JAMA Pediatr. doi:10.1001/jamapediatrics.2018.5056
Published online January 28, 2019.

Supplemental content

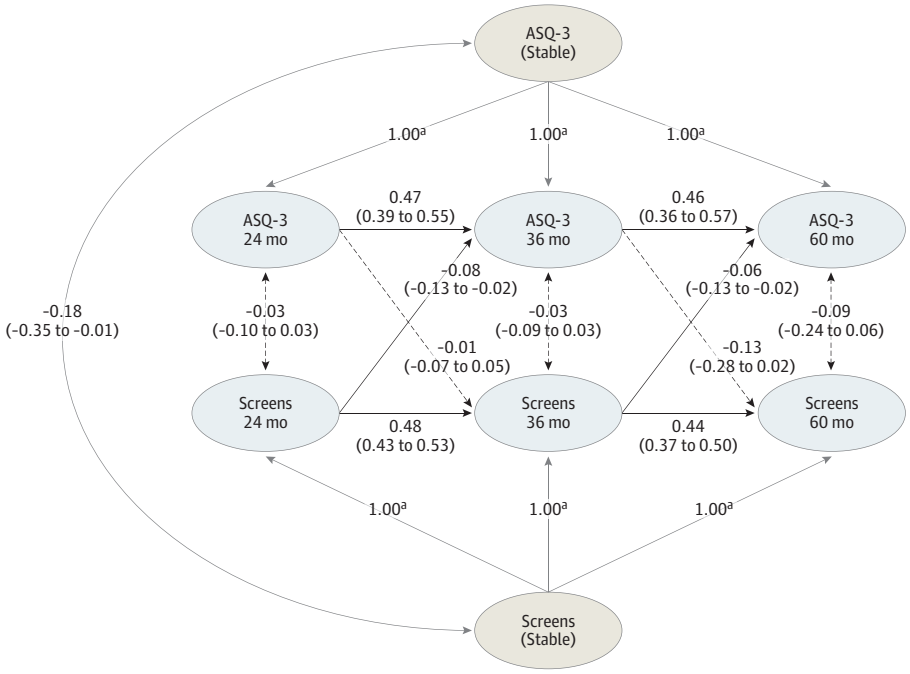
Author Affiliations: Department of Psychology, University of Calgary, Calgary, Alberta, Canada (Madigan, Racine, Mori); Department of Paediatrics, Alberta Children's Hospital Research Institute, Calgary, Alberta, Canada (Madigan, Racine, Mori, Tough); Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada (Browne).

Corresponding Author: Sheri Madigan, PhD, Department of Psychology, University of Calgary, 2500 University Ave, Calgary, AB T2N 1N4, Canada (sheri.madigan@ucalgary.ca).

By school entry, 1 in 4 children shows deficits and delays in developmental outcomes such as language, communication, motor skills, and/or socioemotional health.[1,2] Thus, many children are beginning school inadequately prepared for learning and academic success. Gaps in development tend to widen vs shrink over time without intervention,[3] creating a burden on education and health systems in the form of greater government and public expenditures for remediation and special education.[4,5] Consequently, there have been efforts to identify factors, including children's screen time,[6] that may create or exacerbate disparities in early child development.

Digital media and screens are now ubiquitous in the lives of children. Approximately 98% of US children aged 0 to 8 years live in a home with an internet-connected device and, on average, spend over 2 hours a day on screens.[7] This amount exceeds the recommended pediatric guideline that children spend no more than 1 hour per day viewing high-quality programming.[8,9] Although some benefits of high-quality and interactive screen time have been identified,[10-13] excessive screen time has been associated with a number of deleterious physical, behavioral, and cognitive outcomes.[14-21] While it is possible that screen time interferes with opportunities for learning and growth, it is also possible that children with delays receive more screen time to help modu-

Figure. Random-Intercepts, Cross-Lagged Panel Model Illustrating Within-Person Association Between Developmental Outcomes (Ages and Stages Questionnaire, Third Edition [ASQ-3]) and Screen Time (Hours per Week) from Ages 24 to 60 Months, Controlling for Between-Person Differences



Standardized estimates (95% CIs) are presented. Solid lines represent estimates where 95% CIs do not include zero. The central, blue-tinted part of the model connected by black solid and dashed lines is the within-person (dynamic) part, and the outer, tan-tinted part of the model connected by gray lines is the between-person (stable) component.

[a] Pathways constrained to 1.00 to isolate between-person factor (n = 2441).

age, than other children. In addition, a statistically significant and negative covariance between the between-person components suggests that children with higher levels of screen time exhibit poorer performance on developmental screening tests, on average, and across all study waves.

In the time-variant component of the model, statistically significant autocorrelations for every estimated lag indicate substantial within-person stability in constructs over time. As detailed in the Figure, after accounting for this within-person stability, there were significant and negative cross-lags linking screen time exposure at 24 months with lower scores on developmental screening tests at 36 months (β, −0.08; 95% CI, −0.13 to −0.02), and also with screen time exposure at 36 months associated with lower scores on developmental screening tests at 60 months (β, −0.06; 95% CI, −0.13 to −0.02). The obverse direction of lower scores on developmental screening tests being associated with higher levels of later screen time was not observed. Also, within-time covariances were not significant. Taken together, these findings suggest that higher levels of screen exposure relative to a child's average level of screen time were associated with significantly poorer performance on developmental screening tests at the next study wave relative to a child's average level of developmental milestones but not vice versa.

Between-Person Predictors of Average Screen Time and Developmental Outcomes

Covariates were treated as predictors in a multivariate regression, whereby the between-person factors were regressed onto all variables simultaneously. The forced entry of all of these covariates resulted in a poorer-fitting model, although the permission of a covariance matrix among all covariates yielded a model that fit moderately well on fit indexes, with the exception of the TLI (χ^2_{53} = 521.04; P < .001; RMSEA = 0.06; 95% CI, 0.05-0.06; TLI = 0.78; SRMR = 0.067). As detailed in **Table 2**, higher person-level means on the ASQ-3 were observed for girls and when mothers reported lower maternal depression and higher household income, maternal positivity, levels of child physical activity, child exposure to reading, and hours of sleep per day. These predictors accounted for 15% of the variance. Lower person-level means of screen time were observed for girls and when mothers reported

Table 2. Between-Person Predictors of Average Screen Time and Developmental Milestones

	Standardized Estimate (β), 95% (CI)	
Predictor	Developmental Outcomes	Screen Time
Child age	0.04 (−0.01 to 0.09)	−0.02 (−0.07 to 0.03)
Maternal age	0.00 (−0.06 to 0.07)	0.03 (−0.03 to 0.09)
Female child	0.23 (0.18 to 0.27)[a]	−0.06 (−0.11 to −0.02)[a]
Income	0.11 (0.06 to 0.16)[a]	−0.10 (−0.15 to −0.04)[a]
Educational level	0.03 (−0.02 to 0.08)	−0.19 (−0.25 to −0.14)[a]
Physical activity	0.07 (0.01 to 0.12)[a]	−0.01 (−0.07 to 0.04)
Maternal positivity	0.13 (0.08 to 0.18)[a]	−0.03 (−0.08 to 0.02)
Reading to child	0.12 (0.06 to 0.18)[a]	−0.08 (−0.13 to −0.02)[a]
Maternal depression	−0.06 (−0.11 to −0.01)[a]	0.08 (0.03 to 0.13)[a]
Sleep (h/night)	0.11 (0.06 to 0.16)[a]	−0.14 (−0.19 to −0.10)[a]
Child in care	0.02 (−0.03 to 0.06)	−0.03 (−0.09 to 0.00)
R^2	0.15 (0.12 to 0.19)[a]	0.12 (0.08 to 0.15)[a]

[a] Estimates in which 95% CIs do not include 0.

(1981), explains how misbehaving children have benefitted from these techniques. Step eight (never give up) says, "Always look for a better way." Let the client or student know you care what happens, and try as many plans as you must to establish better behaviors.

4. Reality Therapy Programs in Schools

Glasser's book, *Schools Without Failure,* presents an understandable way to apply these ideas to schools. Published in 1969, it was the largest selling American book on education during the 1970s. Enthusiasm from educators who were using these ideas helped Glasser to found the Educator Training Center in 1968, to teach how to use the concepts in schools.

Starting with educational films and inservice training programs, the Educator Training Center has expanded until the staff now help universities. For example, the University of Wisconsin at LaCrosse offers master's degrees on these concepts. At least 250,000 teachers in the United States have been specifically trained by the Educator Training Center staff and countless others have taken courses by those who have had this training. In translating reality therapy to schools, educators have recognized that the most important concepts to be learned are: (a) to base instructional programs on success practices—increase the student's success within the curriculum; (b) use the classroom meeting as a device for involvement between teacher and student, and student and student. The class and teacher sit in a circle to communicate in a nonjudgmental way and discuss values, goals of the class, how to live together successfully in the classroom, and anything else that can increase involvement and help students to know someone cares. Use of these meetings shows significant improvement in behavior and increase in learning. This process is best described in the book *Schools Without Failure* (1969). In *Focus on Guidance*, Thompson and Cates (1976) quote a study in Tennessee showing specific data for frequency of inappropriate behavior and how it decreases as the classroom teacher uses these methods. Along with this comes increased learning.

The Educator Training Center has developed programs which help school personnel discipline students and involve them in their learning in a way that helps them to accept responsibility for all their behavior. A description of how this can be done can be studied in a pamphlet called "*Glasser's Approach to Discipline*" (Educator Training Center 1977). It is a 10-step program. The first 3 steps use the process of reality therapy to look at how one deals with disruptive students. Steps 4 to 7 are specific ways to deal with children having difficulty, and the last 3 steps bring in other resources within the school and/or community. Schools which have applied this program report decreases in suspensions by 50–80 percent in junior and senior high schools, and vandalism by 40–90 percent. Improvement in teacher morale and professional growth were also cited as being significant gains by schools using this program. Administrators gained time to be the educational leaders they were meant to be. The implementation of this program can be further understood by reading a book by a principal, Bill Borgers, *Return to Discipline* (1979), who applied it to the students, teachers, and counselors in his school. He found that by having everyone understand how to help students take responsibility for their own behaviors, schools could function in a way that produced more learning.

Another book, *Loneliness in the Schools* by Marc Roberts (1973) shows how to use the schools-without-failure principles to establish more effective interaction within schools. His plan to encourage humaneness and growth within the school environment was created out of his concern for the loneliness of people who spend many hours each day in school.

Although the Educator Training Center works throughout the United States, its main office is in Long Beach, California.

See also: Counseling Theories

Bibliography

Borgers W 1979 *Return to Discipline.* Privately printed, Austin, Texas

Educator Training Center 1977 *Glasser's Approach to Discipline.* Long Beach, California

Glasser N 1980 *What Are You Doing? How People are Helped Through Reality Therapy: Cases.* Harper and Row, New York

Glasser W 1965 *Reality Therapy: A New Approach to Psychiatry.* Harper and Row, New York

Glasser W 1969 *Schools Without Failure.* Harper and Row, New York

Glasser W 1981 *Stations of the Mind: New Directions for Reality Therapy.* Harper and Row, New York

Medick J M 1981 *Classroom Misbehavior: Turning It Around.* Fanning, East Lansing, Michigan

Roberts M 1973 *Loneliness in the Schools.* Argus Communications, Niles, Illinois

Thompson C L, Cates J T 1976 Teaching Discipline to Students: An Individualized Teaching–Counseling Approach. *Focus on Guid.* 9(4)

W. Glasser: N. Glasser

Reciprocal Effects, Analysis of

The analysis of reciprocal effects consists of the study of mutual influences between two (or more) variables. Often, researchers attribute causality to the influences among variables and speak of reciprocal causal effects. Examples from educational research include analyses of the reciprocal effects between: teacher expectations and student achievement, student self-concept and academic achieve-

ment, and vocabulary and comprehension skills. Also, among the many studies of reciprocal effects in psychology are analyses of mother–child interactions, infant intelligence and behavior, and aggressive behavior and viewing of violent television shows. Often, research questions about reciprocal effects have been simplified to, does X cause Y or does Y cause X? Consequently, many researchers have sought only determinations of causal predominance or of *the* causal ordering of the variables.

Longitudinal data are crucial to the analysis of reciprocal effects, because the temporal ordering is needed to unravel the influences linking the variables. Typically, observations on a large number of cases are obtained for each variable on a few (two or more) occasions. (Each case may be an individual or a unit such as a student–teacher or mother–child dyad.) Much of the empirical research and methodological discussion on reciprocal effects has been limited to the two-wave, two variable (2W2V) panel design, in which measures of X and Y are available on each of two occasions

Four different approaches and statistical methods for quantitative data have been used in analyses of reciprocal effects. These methods are cross-lagged correlation, structural regression models, continuous-time feedback models, and multiple time series. Separate methods for dichotomous and categorical data, which have not been much used in educational research, are Lazarsfeld's 16-fold table, log-linear models for contingency tables, and continuous-time Markov models (see *Structural Equation Models*; *Log-linear Models*; *Contingency Tables*).

1. Cross-lagged Correlation

Cross-lagged correlation has been the most popular procedure in educational and psychological research for the analysis of reciprocal effects. Most often, cross-lagged correlation is used to determine a predominant causal influence. Figure 1, which presents the population correlations among the variables in a 2W2V panel design, is the diagram that accompanies expositions of cross-lagged correlation. The population cross-lagged correlations are $\rho_{X_1 Y_2}$ and $\rho_{Y_1 X_2}$.

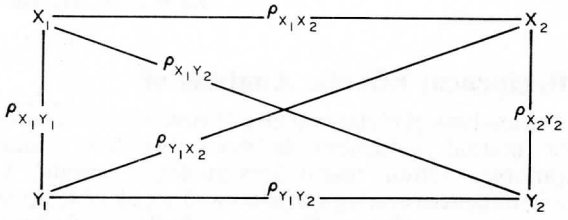


Figure 1
Population correlations for a 2W2V panel. The variables at time 1 are X_1 and Y_1 and the variables at time 2 are X_2 and Y_2

The difference between the population cross-lagged correlations, $\rho_{X_1 Y_2} - \rho_{Y_1 X_2}$, is the basis for attributions of a predominant causal influence. If the data indicate that $\rho_{X_1 Y_2} - \rho_{Y_1 X_2}$ is positive, the predominant causal influence is concluded to be in the direction of X causing Y. If the data indicate that $\rho_{X_1 Y_2} - \rho_{Y_1 X_2}$ is negative, the predominant causal influence is concluded to be in the direction of Y causing X. Usually, attributions of predominant causal influences are made only when the null hypothesis of equal population cross-lagged correlations is rejected. If this null hypothesis is not rejected, the usual interpretation is that no direct causal influences exist between X and Y; in particular that a common causal influence may be responsible for their observed associations. This interpretation has been adopted as a null hypothesis of spuriousness, which is represented through a model allowing no direct influences between X and Y but with an unmeasured third variable influencing both X and Y at each time (Kenny 1979 Fig. 12-2, Rogosa 1980 Fig. 3).

The extension of cross-lagged correlation to determinations of causal predominance when more than two waves of data are available is to compare the cross-lagged correlations from all possible two-wave combinations. Rogosa (1980) showed that this strategy of using the multiple waves for replication of differences between the cross-lagged correlations is more likely to generate confusion than corroboration.

A related statistical procedure for 2W2V panel data, which has seen a number of applications in educational research, is the frequency-in-change-in-product–moment procedure developed by Yee and Gage (1968). Although the Yee–Gage procedure differs from cross-lagged correlation in many important details, this procedure for the analysis of reciprocal effects suffers from the same basic deficiencies as cross-lagged correlation.

Cross-lagged correlation does not provide dependable information as to the causal structure underlying the data. Building upon earlier analyses by Duncan (1969) and Heise (1970), Rogosa (1980) demonstrated that, when there are no reciprocal causal effects, the difference between the cross-lagged correlations may be small or may be large; and when there are considerable reciprocal causal effects, the difference between the cross-lagged correlations may be small or may be large. Furthermore, a zero difference between the cross-lagged correlations (indicating spuriousness in cross-lagged correlation) is consistent with large reciprocal causal influences or with small or nonexistent causal influences between the variables. Moreover, cross-lagged correlation may indicate a causal predominance opposite to that of the actual causal structure of the data. Hence, neither determinations of spuriousness nor causal predominance can be trusted.

A basic deficiency in cross-lagged correlation is

the lack of an explicit definition of a causal effect. Without a clearly defined quantity to be estimated, it is not surprising that cross-lagged correlation fails to provide sound inferences. Also, the emphasis on causal predominance in cross-lagged correlation is unwise. The reciprocal nature of many social and educational processes makes determination of only causal predominance a serious oversimplification of the research problem. Measures of the strength and duration of the reciprocal relationship and of the specific causal effects are needed.

2. Structural Equation Models

Structural regression formulations of reciprocal effects in longitudinal panel data were originally introduced in the path analysis literature (Wright 1960, Duncan 1969, Heise 1970). The term causal model is popular for describing both the path analysis and the more general structural equation models. In these models, a causal effect is represented by the change in an outcome variable that results from an increment to an antecedent variable. For two variables, X and Y, the reciprocal influences are represented by the regression parameters of the path from a prior X to a later Y and from a prior Y to a later X.

Previous formulations of regression models for panel data with reciprocal causal effects have focused on models for 2W2V data. Figure 2 represents a specific regression model for 2W2V data, given by the regression equations:

$$X_2 = \beta_0 + \beta_1 X_1 + \gamma_2 Y_1 + u \quad (1)$$

$$Y_2 = \gamma_0 + \beta_2 X_1 + \gamma_1 Y_1 + v \quad (2)$$

The parameters β_2 and γ_2 represent the lagged, reciprocal causal effects between X and Y and thus are of central importance in the investigation of reciprocal causal effects.

Although the structural regressions do provide a model that defines reciprocal effects among the variables, the validity of inferences about the reciprocal effects depends crucially on the validity of the model. Foremost among the important assumptions built into Fig. 2 and Eqns. (1) and (2) is that X and Y constitute a closed system so that no important influences have been omitted from the regression model. Also important is the assumption that all causal influences are lagged; simultaneous causal influences between X and Y are not included. With only 2W2V data, frequently it is not possible to distinguish between different underlying models, which makes the determination of reciprocal effects very difficult. Additional observations can aid the formulation and testing of the regression models; one example of the use of regression models for three waves of data is the analysis of the influences between economic development and educational expansion in Hannan et al. (1974).

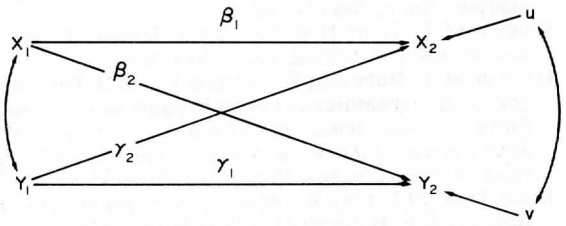


Figure 2
Representation of a regression (path analysis) model for 2W2V data

A generalization of these path analysis models is to specify X and Y to be latent variables having multiple indicators at each time point. The influences among the variables are represented by the parameters of the structural regression equations that relate the latent variables. (A measurement model connects the latent variables with their indicators.) For example, recasting Eqns. (1) and (2) in terms of latent variables, β_2 and γ_2 would then represent the reciprocal effects between the latent variables X and Y. Examples of the use of structural equation models for the analysis of reciprocal effects are: the analysis of attitudes and behaviors in Bentler and Speckart (1981), the analysis of intellectual flexibility and complexity of work in Kohn and Schooler (1978), and models for home environment and intellectual development in Rogosa (1979).

3. Continuous-time Feedback Models

An alternative formulation of reciprocal effects is to model the rates of change of the variables. A simple two variable continuous-time model that incorporates reciprocal influences between X and Y is:

$$\frac{dX(t)}{dt} = b_0 + b_1 X(t) + c_2 Y(t) \quad (3)$$

$$\frac{dY(t)}{dt} = c_0 + c_1 Y(t) + b_2 X(t) \quad (4)$$

Equations (3) and (4) are coupled differential equations which stipulate that the rates of change of X and Y at any time depend linearly on the levels of X and Y. The parameters b_2 and c_2 represent the cross effects or couplings between X and Y. Note that Eqns. (3) and (4) are deterministic. Similar models for change can be formulated which include stochastic components, exogenous variables, and other generalizations. Many applications of these models are presented in Hannan and Tuma (1983).

Rates of change are not directly observable. However, the solution of the system of differential equations in Eqns. (3) and (4) yields equations, in terms of the observable variables, of the same form as Eqns. (1) and (2). The parameters β_2 and γ_2 in

Eqns. (1) and (2) are nonlinear functions of the time between observations and the parameters of Eqns. (3) and (4). That the solutions of Eqns. (3) and (4) have the same form as the regression model for 2W2V data allows models for 2W2V data to be thought of as reflecting a more general process, that of causal influences and resulting adjustments that are continuous in time.

4. Multiple Time Series

The statistical analysis of reciprocal effects, or feedback, between two time series has been an active area in econometrics, with most applications investigating reciprocal effects between money supply and income or those between advertising and sales. The statistical methods are based on predictability criteria. Loosely speaking, one time series, $X(t)$, is said to cause another time series, $Y(t)$, if present Y values can be predicted better using past values of X than by not using past values of X, other relevant information (including past values of Y) being included in the prediction. This definition encompasses both lagged and instantaneous influences between X and Y. A comprehensive classification of the possible patterns of causal influences was provided by Pierce and Haugh (1977) who also presented data analysis procedures based on the correlations of residuals between the separately filtered time series for detecting these reciprocal effects. In addition, useful measures of linear dependence and feedback among multiple time series were developed by Geweke (1982).

In most economic research on reciprocal effects, the data consist of a single extensive time series of observations for each variable. The minimum number of observations over time needed for the application of time series statistical models is far beyond the usual design of longitudinal research in education. The longitudinal data for which methods commonly used in education (cross-lagged correlation, structural regression, and continuous-time feedback models) are applicable consist of a collection of many replications of very short time series (often with only two observations). Of course, such limited temporal data cannot support the complex time-series models used in the econometric analyses of reciprocal effects. One psychological application of models and analyses of reciprocal effects in time series data is the analysis of play behavior of individual mother–infant dyads in Gottman and Ringland (1981).

5. Methods for Categorical Data

Analyses of reciprocal effects using dichotomous or polychotomous variables have not been common in educational research. Lazarsfeld's 16-fold table for analyzing reciprocal effects among dichotomous variables is the best-known method for categorical data; Lazarsfeld (1978) provided a history of the development and application of this procedure. A natural extension of the 16-fold table analysis is the application of log-linear models for contingency tables (Goodman 1973). A third approach to the analysis of reciprocal effects is the use of continuous-time stochastic models, in particular, discrete-state, continuous-time Markov models. For additional references and for applications of these methods, see Coleman (1968), Hannan and Tuma (1979, 1983), and Markus (1979).

6. Conclusion

The investigation of reciprocal effects is an extremely difficult enterprise. Questions about reciprocal effects are some of the most complex in the design and analysis of longitudinal research. A humbling reality for research on reciprocal effects is that much simpler, preliminary research questions, namely those connected with the measurement of individual change and the assessment of correlates of change, remain controversial and unsolved. Clearly, reciprocal effects cannot be studied cheaply. Extensive, high-quality longitudinal data and theoretically based, explicit models of the reciprocal effects are absolutely necessary.

Bibliography

Bentler P M, Speckart G 1981 Attitudes "cause" behaviors: A structural equation analysis. *J. Pers. Soc. Psychol.* 40: 226–38

Coleman J S 1968 The mathematical study of change. In: Blalock H M, Blalock A B (eds.) 1968 *Methodology in Social Research.* McGraw-Hill, New York, pp. 428–78

Duncan O D 1969 Some linear models for two-wave, two-variable panel analysis. *Psychol. Bull.* 72: 177–82

Geweke J 1982 The measurement of linear dependence and feedback between multiple time series. *J. Am. Stat. Ass.* 77: 304–24

Goodman L A 1973 Causal analysis of data from panel studies and other kinds of surveys. *Am. J. Sociol.* 78: 1135–91

Gottman J M, Ringland J T 1981 The analysis of dominance and bidirectionality in social development. *Child Dev.* 52: 393–412

Hannan M T, Tuma N B 1979 Methods for temporal analysis. *Annu. Rev. Sociol.* 5

Hannan M T, Tuma N B 1983 *Social Dynamics: Models and Methods.* Academic Press, New York

Hannan M T, Rubinson R, Warren J T 1974 The causal approach to measurement error in panel analysis: Some further contingencies. In: Blalock H M Jr (ed.) 1974 *Measurement in the Social Sciences: Theories and Strategies.* Aldine, Chicago, Illinois, pp. 293–323

Heise D R 1970 Causal inference from panel data. In: Borgatta E F, Bohrnstedt G W (eds.) 1970 *Sociological Methodology.* Jossey-Bass, San Francisco, California

Kenny D A 1979 *Correlation and Causality.* Wiley, New York

Kessler R C, Greenberg D F 1981 *Linear Panel Analysis: Models of Quantitative Change.* Academic Press, New York

Kohn M L, Schooler C 1978 The reciprocal effects of the substantive complexity of work and intellectual flexibility: A longitudinal assessment. *Am. J. Sociol.* 84: 24–52

Lazarsfeld P F 1978 Some episodes in the history of panel analysis. In: Kandel D B (ed.) 1978 *Longitudinal Research on Drug Use.* Wiley, New York, pp. 249–65

Markus G B 1979 *Analyzing Panel Data.* Sage, Beverly Hills, California

Pierce D A, Haugh L D 1977 Causality in temporal systems: Characterizations and a survey. *J. Econometrics* 5: 265–93

Rogosa D R 1979 Causal models in longitudinal research: Rationale, formulation, and interpretation. In: Nesselroade J R, Baltes P B (eds.) 1979 *Longitudinal Methodology in the Study of Behavior and Development.* Academic Press, New York, pp. 263–302

Rogosa D R 1980 A critique of cross-lagged correlation. *Psychol. Bull.* 88: 245–58

Wright S 1960 The treatment of reciprocal interaction, with or without lag, in path analysis. *Biometrics* 16: 423–45

Yee A H, Gage N L 1968 Techniques for estimating the source and direction of causal influence in panel data. *Psychol. Bull.* 70: 115–26

D. R. Rogosa