

Simultaneous Feedback Models with Macro-Comparative Cross-Sectional Data

Breznau, Nate

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Breznau, N. (2018). Simultaneous Feedback Models with Macro-Comparative Cross-Sectional Data. *Methods, data, analyses : a journal for quantitative methods and survey methodology (mda)*, 12(2), 265-307. <https://doi.org/10.12758/mda.2017.07>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by/3.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more Information see:
<https://creativecommons.org/licenses/by/3.0>

Simultaneous Feedback Models with Macro-Comparative Cross-Sectional Data

Nate Breznau

Mannheim Centre for European Social Research (MZES)

Abstract

Social scientists often work with theories of reciprocal causality. Sometimes theories suggest that reciprocal causes work simultaneously, or work on a time-scale small enough to make them appear simultaneous. Researchers may employ simultaneous feedback models to investigate such theories, although the practice is rare in cross-sectional survey research. This paper discusses the certain conditions that make these models possible if not desirable using such data. This methodological excursus covers the construction of simultaneous feedback models using a structural equation modeling perspective. This allows the researcher to test if a simultaneous feedback theory fits survey data, test competing hypotheses and engage in macro-comparisons. This paper presents methods in a manner and language amenable to the practicing social scientist who is not a statistician or matrix mathematician. It demonstrates how to run models using three popular software programs (*MPlus*, *Stata* and *R*), and an empirical example using *International Social Survey Program* data.

Keywords: simultaneous feedback model, cross-sectional data, macro-comparative research, structural equation modeling, reciprocal causality, *Mplus*, *Stata*, *R* (lavaan)



© The Author(s) 2018. This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 License. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Social scientists often study reciprocally causal phenomena. For example, supply and demand in economics; candidate evaluations and party identification in political science; road investment and travel demand in geography; and educational attainment and parenthood entry in sociology and demography (Marini, 1984; Page & Jones, 1979; Xie & Levinson, 2010). When timings of reciprocal causes are unobservable or occur contemporaneously, a state of simultaneous feedback exists. Rather than in cycles, events happen at the same time. Philosophers of causality question the existence of simultaneous feedback (Mulaik, 2009: Chapter 3); however, researchers regularly face theoretical and data conditions that force them to accept simultaneous feedback in practice. This is particularly acute in macro-comparative survey research where observations take place over a year, but theoretical causes may take place at less-than-yearly intervals. All sub-yearly causal effects appear simultaneous within a year interval. Under certain conditions, macro-comparative researchers can employ simultaneous feedback models (SFMs) to capture these effects, allowing them to overcome some limitations of comparative cross-sectional survey research.

Herein, I elaborate when and how to use SFMs. This requires structural equation modeling (SEM) strategies to explicate theoretical relationships before extracting meaningful statistical results. I use minimal statistical and mathematical jargon without matrix algebra¹, and a practical example of public opinion and social policy. I show that SFMs provide a powerful method for macro-comparative survey researchers to explain, predict and compare reciprocally causal phenomena.

Simultaneous Feedback

Instances where two phenomena are co-causes of each other are ubiquitous in social research²; however, modeling reciprocal causality is challenging. Time is usually

-
- 1 Matrix algebra is the basis of nearly all social science statistics including SFMs; however, this excursus is for the practicing social scientist who is unlikely a matrix algebraician.
 - 2 More non-exhaustive examples: (Brehm & Rahn, 1997; Chong & Gradstein, 2007; Claibourn & Martin, 2000; Liska & Reed, 1985; Mulatu & Schooler, 2002; Owens, 1994; Thornton, Axinn, & Hill, 1992)

Acknowledgments

I am grateful to two anonymous reviewers, discussions on SEMNET, Bart Meuleman and Sebastian Pink for helpful instruction and comments.

Direct correspondence to

Nate Breznau, Mannheim Centre for European Social Research (MZES)
E-Mail: breznau.nate@gmail.com

the basis for explaining or predicting things (Elwert, 2013; Pedhazur, 1997). To be a cause or a useful predictor, X must take place prior to Y . If X happened after Y it is not a cause³. Sometimes researchers cannot effectively observe or operationalize time. For example, the moods of roommates are theoretically timed causes of each other but may unfold so quickly that they appear simultaneously causal (Siegel & Alloy, 1990). It is possible that there are nanoseconds in between, but these are unobservable. Furthermore, excessive complexity of timings and multitudinous mood causes running in both directions leave the researcher viewing mood effects as simultaneous.

Macro-comparative research is similar on a larger time scale. Contextual data tend to measure time points spanning an entire year. Reciprocally causal effects that take place in just days, weeks or even months subsume into these yearly observations. For example, public opinion likely causes changes in policymaking on a weekly or monthly basis as policymakers constantly try to meet public preferences. Simultaneously, public opinion changes within minutes or hours in response to policy changes. When capturing these opinion-policy effects with survey data, the two appear to have simultaneous causality within each year unit. Moreover, survey researchers lack yearly comparative opinion data across countries, e.g., cross-sectional yearly time-series⁴, rendering longitudinal methods sometimes inappropriate. Having sporadic macro-comparative survey data means SFMs might be appropriate, but this is not a sufficient condition to use them. Theory must drive this decision (Hayduk et al., 2007; Kaplan, Harik, & Hotchkiss, 2001).

Given a theory of simultaneous feedback between two phenomena, I label them Y_1 and Y_2 ⁵, where at least two different linkages exist between them if not more. One for the effect of Y_1 on Y_2 and one vice-versa. However, when I observe and quantify Y_1 and Y_2 as variables, they have only one empirical linkage: their covariance (or correlation). Identifying two effects statistically, when there is only one covariance, is not possible. Y_1 and Y_2 are *nonrecursive* meaning that their respective effects on each other cannot be identified using only their joint information. Their reciprocal relationship makes them *endogenous* meaning caused from

3 The method herein applies to causal or explanatory research subsuming causes or several causes into a package of predictive power without considering the mechanisms in detail. Although causality is at the heart of the theoretical side of SFMs, the vast realm of mathematics and philosophy of causality is beyond the scope of this paper (Pearl, 2010; Sobel, 1996).

4 Although impressive, many macro-comparative sources of survey data barely qualify as longitudinal, cross-sectional time-series when fielded only every 2 to 10 years (e.g., *European Social Survey*, *World Values Survey* and *International Social Survey Program*).

5 I use Y_1 and Y_2 rather than X and Y , because Y denotes dependent variables. Reciprocally causal variables are dependent on each other.

within; however, identifying these nonrecursive endogenous effects requires some *exogenous* causes from without.

I describe this problem using Equations 1 and 2, and Figure 1. Both cases present a system logically *underidentified* – there are more parameters to be estimated than pieces of observed information (two coefficients b_1 and b_2 yet only one covariance of Y_1 and Y_2).

$$Y_1 = b_1 Y_2 + e_1 \quad (1)$$

$$Y_2 = b_2 Y_1 + e_2 \quad (2)$$

Regression analysis could estimate Equations 1 and 2, but results are probably inaccurate given a theory of reciprocal causality. In Figure 1 the arrows represent theoretical effects, and b_1 and b_2 represent regression coefficients. Y_1 is not known without knowing Y_2 and Y_2 is not known without knowing Y_1 : An endless circle!

Identifying b_1 and b_2 is an exercise in finding more variables or parameters. Figure 2 gives four common formal models containing reciprocal causality, some identified, others not. Adding *instrumental variables* (IVs) enables identification of unique b_1 and b_2 effects. An *IV* is *exogenous*: not caused by the system described in the model, not caused by Y_1 or Y_2 and not moderating or somehow causing the causal paths linking Y_1 and Y_2 . Figure 2A describes some phenomenon labeled Y_1 occurring at time “t” that is both a cause (arrow pointing away) and outcome (arrow pointing towards) of another phenomenon Y_2 measured at the same time. In this, IV_1 must be a cause of Y_1 but not of Y_2 ; and IV_2 must cause Y_2 but not Y_1 (see section “Instrumental variables”).

Figure 2A is the basic SFM form.

Other common reciprocal effects models appear in Figure 2B-2D. Cross-lagged reciprocal effects (2B) are a common form of reciprocal causal modeling (for discussions: Billings & Wroten, 1978; Schaubroeck, 1990). Looking at Y_1 and Y_2 longitudinally over time generates separate, unique covariances between Y_1 and Y_2 ; one for $Y_{2,t-1}$ with $Y_{1,t}$ and another for $Y_{1,t-1}$ with $Y_{2,t}$. Cross-lagged models require the assumption that Y_1 and Y_2 do not cause each other simultaneously for identification (omitted arrows between them at time t). Macro-comparative survey researchers rarely have sequential time series of survey data in several countries making these models untenable, often because of missing time points or the exact

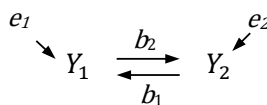


Figure 1 Path Model of Equations 1 and 2

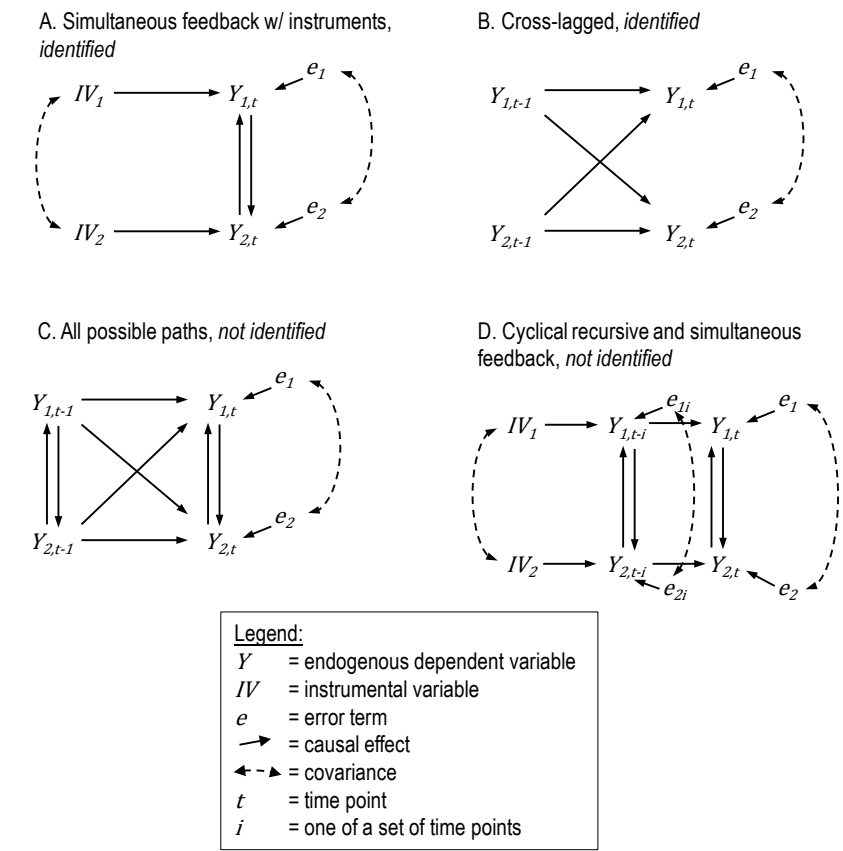


Figure 2 Various Models of Reciprocal Causality

timing of cause and effect do not match the starting and ending points of the survey (Finkel, 1995). If causes occur at a less-than-yearly interval, in addition to across time-units, then Figure 2C is accurate visually but underidentified statistically. A similar story occurs when adding instrumental variables to 2C as shown in 2D. The instruments do not add enough power to overcome the cyclically recursive problem of observing Y_1 and Y_2 over time because they are causes of their later selves in addition to causing each other leading again to too many parameters.

Conditions Necessary for Simultaneous Feedback Models

A strong theory, equilibrium, model identification and appropriate instrumental variables are the necessary features to employ Figure 2A.

Theory

The first and most important requirements of SFMs are theoretical. Without theory, the two arrows connecting Y_1 and Y_2 do not exist. There must be an *a priori* logic to the data-generating model, defensible against confounding effects (Heckman, 2000; Rigdon, 1995). Thus, *a theory of simultaneous causality is the baseline condition*. This theory must specify that during the observational window causal effects materialized between Y_1 and Y_2 ; regardless of whether these are direct or operating through intermediary mechanisms. A researcher must provide sufficient argument for simultaneity. That of, (1) *co-determinacy* with effects that happen ‘instantaneously’ in less time than can be observed, or (2) *complexity* with effects that are constantly taking place going in many directions having various lengths of time to complete; so as to appear simultaneous. Without this theoretical basis to the Y_1 and Y_2 relationship, researchers have no ground to stand on in defense of simultaneous feedback (Hayduk et al., 2007; Markus, 2010). Theory determines the design of a formal path model, instrumental variables, equilibrium, size and direction of effects, the set of independent variables, and the nature of errors and estimation techniques. Suffice to say, theory is paramount.

Equilibrium

Two forms of equilibrium need be present in SFMs. The first is that causal effects are theoretically stable or behave in a stable manner. There should be logical argument that the impact of Y_1 on Y_2 and vice-versa, do not change over time (Kaplan et al., 2001). In other words, the effects should not depend on when in time the researcher observes Y_1 and Y_2 (Sobel, 1990). This is a grey area as inevitably all social things change over time; so a better stance to defend might be they do not change much in a given period. For example, if the area of farmed land reduces the hunger in a society while the rate of hunger increases the area of farmed land, a researcher might argue for equilibrium, as a change in one produces a predictable change in the other. Statistically speaking the regression coefficients should be stable. However, technology increases food produced per acre, disrupting the equilibrium because each acre has a larger impact on hunger reduction. This implies that the regression coefficients change if technology changes, but might be stable

before and after. If the model includes events before and after this change, it is misspecified as a SFM.

The second part is that the causal effects are part of a context at equilibrium, e.g., a political or judicial system. If a system experiences shocks then equilibrium is unlikely, e.g., disruptive wars or economic recessions. Therefore, the researcher must rule out changes to the larger systems within which Y_1 and Y_2 operate (see section “Disequilibrium”).

Identification

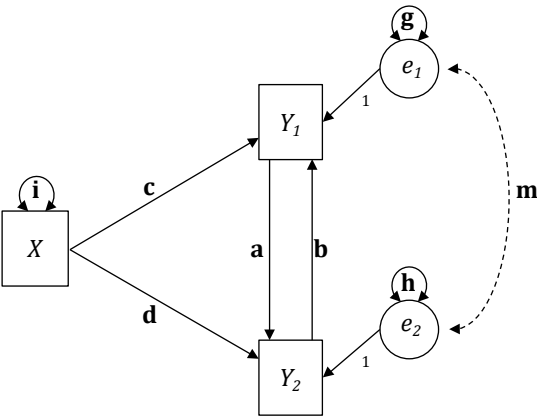
Any formal model, including one with simultaneous feedback must be identified to produce meaningful results or results at all⁶. To identify two statistical coefficients that capture two theoretical effects between Y_1 and Y_2 there must be more than one covariance in the model. Only one covariance in the feedback model is *underidentified*, meaning more parameters to estimate than pieces of observed information leading to a negative value for *model degrees of freedom*. Pieces of observed information are all parameters the researcher observes in the data including the means, variances and covariances of the variables in the model, also known as model “elements” (Rigdon, 1994). In SFMs, the observed means are often not estimated because researchers’ main interests are in the coefficients between Y_1 and Y_2 that derive entirely from covariances, irrespective of means. Adding means to the analysis generally complicates things with few cases.

Without means, the formula to calculate pieces of observed model information is $v(v+1)/2$, where v is the number of observed variables (Kline, 2011). The model needs a minimum of the same number of model elements as freely estimated parameters for identification, i.e., model degrees of freedom needs to be larger than or equal to zero. To illustrate, I add one predictor variable X , as shown in Figure 3. Figure 3A is not identified because it requires estimation of four coefficients (**a** through **d**) and three variances (**g** through **i**), with residual covariance **m** optional. Fixing **m** to zero for now, and knowing nothing about **a** through **i**, there are seven freely estimated parameters (**a** through **i**). That means I need seven pieces of information for a just-identified model. There are only six pieces in Figure 3A: three covariances ($X, Y_1 | X, Y_2 | Y_1, Y_2$) and three variances (for X, Y_1 & Y_2), or $3(4)/2=6$. Thus, model degrees of freedom is smaller than zero (six minus seven). Figure 3A is underidentified.

Figure 3B includes IV_1 and IV_2 , creating $5(6)/2 = 15$ pieces of information. Assuming that the IV s and the error terms are correlated (parameters **n** and **m** respectively), the model has 15 freely estimated parameters (all letters in 3B),

6 Any introductory text on structural equation modeling covers identification. I find Kline (2011) a useful source.

A. Without instruments (not identified)



B. With instruments (identified)

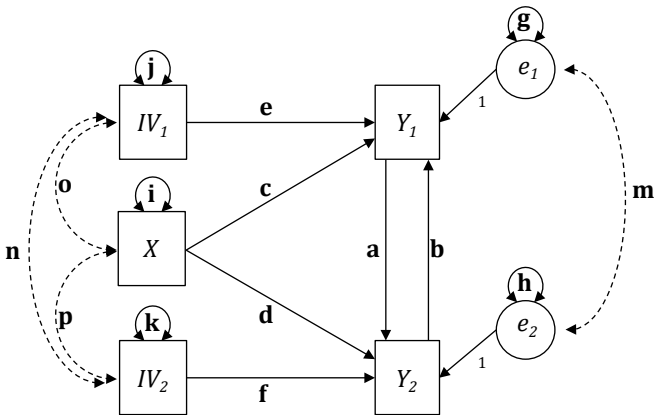


Figure 3 Identifying Simultaneous Feedback Models

meaning model degrees of freedom is zero and the model is just-identified. An ideal model has more than zero, for example three IV s leads to $6(7)/2=21$ pieces of observed information and 20 freely estimated parameters; degrees of model freedom equals one. However, IV s are difficult to find. An identification rule requires at least one IV for each Y variable. If both instruments are attached to Y_2 , and none to Y_1 , the model might have degrees of freedom greater than zero, but the model is still not identified without an IV for Y_1 . This is known as the *rank condition*. This condition is satisfied when, “each variable in a feedback loop has a unique pattern of direct effects on it from variables outside the loop” (Kline, 2011, p. 135). Adding

more X variables to 3B does not help with identification as it does not change the degrees of model freedom nor add *unique* direct effects.

Instrumental Variables

Identification depends on instrumental variables (IV_1 and IV_2). Necessary conditions for selecting IV s are theoretical and statistical. “Instrumental variables” is both an estimation technique and a label for specific exogenous variables (Sargan, 1958). This section is devoted to exogenous *variables*, saying nothing of estimation techniques⁷. An IV must be exogenous to the dependent variable. In experimental language, IV causes the distribution of a treatment but not the outcome. In non-experimental language, the endogenous variable depends on the values of the IV independently from the dependent variable, or the dependent variable only shows covariance with the IV after conditioning on the endogenous variable.

In Figure 3B, the IV for Y_1 must not cause Y_2 . If IV_1 is a cause of Y_2 then IV_1 is an independent variable, not an IV . All independent variables explain or predict all endogenous variables, thus are part of the data-generating model of Y_2 (and Y_1). For IV_1 to pass it must not be part of the data-generating model of Y_2 . This is the *exclusion restriction*. The problem is not correlation of IV_1 with Y_2 , but correlation of IV_1 with e_2 ; i.e., correlation with the unexplained disturbance or error in the dependent variable after adjusting for the impact of all independent variables. If IV_1 causes Y_2 , or omitted variables cause both IV_1 and Y_2 then a correlation of IV_1 with e_2 exists; and the larger this correlation, the larger the problems with the IV . If IV_1 has a small correlation with e_2 because of measurement or random error, then as the sample size approaches infinity the correlation approaches its true value of zero (i.e., asymptotic correlation = 0). If so, small IV_1 with Y_2 correlations after adjusting for covariates are acceptable.

When meeting these conditions, IV_1 and IV_2 decompose the single correlation between Y_1 and Y_2 in Figure 3B into 3 parts: (1) the part that could result from a causal effect or a shared omitted causal effect of Y_1 on Y_2 (covariance left after removing that predicted by IV_2), (2) the same for Y_2 on Y_1 , and (3) the unexplained remaining covariance of error terms e_1 and e_2 . Although technically optional, Part (3) is usually modeled, because finding instruments that explain *everything* about Y_1 and Y_2 with no remainder is unlikely. Moreover, the error term e_1 is produced by a causal effect of Y_2 (path **b**). Yet e_2 is a part of Y_2 and is therefore *by definition* a part of the error term e_1 , i.e., correlated with its own partial correlation produced from Y_2 being regressed on Y_1 (Wong & Law, 1999, p. 73). The same is true for e_2 , and therefore specifying no residual correlation may deny the causally defined

7 Other literature covers this in-depth (Angrist, Imbens, & Rubin, 1996; Angrist & Krueger, 2001; Basile, 2008; Bollen, 2012).

model its own properties. Thus, sometimes a cross-sectional nonrecursive model with correlated errors is the ‘best available’ approximation of cross-lagged reciprocal effects when they are otherwise underidentified.

Even if theoretically not causal, a large correlation between IV_1 and Y_2 is a problem statistically. The larger the correlation the more variance that all independent variables must explain in Y_2 before IV_1 is left uncorrelated with e_2 . In other words, the partial correlation of IV_1 and Y_2 takes away variance in IV_1 that is necessary to explain Y_1 . Thus, the larger this correlation, the greater the disruption of the researcher’s goal to explain variance in Y_1 independent of Y_2 and all independent variables. An inverse of this problem occurs when IV_1 has an increasingly closer-to-zero correlation with Y_1 (Bartels, 1991). The smaller the correlation, the less unique variance of Y_1 that can be explained by IV_1 . These two conditions describe a *weak instrument* problem. Theoretical arguments establish exclusion restrictions necessary to use instrumental variables; however, statistics help identify potential weak instrument problems.

In SEM, model diagnostics, in particular modification indices provide a simple first line of defense to identify weak instruments (see “Fit testing and diagnostics”). This applies because the structural model (what the researcher draws in a path diagram and then programs into the statistical software) fixes the correlation of each IV with each corresponding e to be zero. The fit and modification indices tell the researcher if these fixed zero correlations are realistic given the data. Alternatively, traditional weak instrument tests come from estimating whether results from the instrumental variable estimator and the OLS estimator are consistent, defined in a number of ways depending on the test (Bollen, 2012; Hahn & Hausman, 2002).

There are a variety of statisticians arguing for statistical methods to identify instrumental variables without theoretical arguments that an IV meets the exclusion restriction (see “Other concerns”). Although these methods may asymptotically recover a *known* causal effect (as shown in simulations), the SFM researcher is searching for causal effects whose existence or size is empirically *unknown*. If already known, research becomes unnecessary. Moreover, even when the correlation of IV_1 and Y_2 is exactly zero, there is no statistical way to know for sure that IV_1 and e_2 do not correlate due to causal or omitted variable linkages. Suppression or omitted variables can easily produce a statistical relationship of zero, when the actual causal relationship is non-zero (MacKinnon, Krull, & Lockwood, 2000)⁸. Thus, theoretical arguments are necessary to rule out ‘backdoor’ or confounding relationships among variables. Finally, arguments must establish that the

8 The drawing of a causal structure with a path diagram or graph notation introduced by Wright (1920) allows researchers to follow rules determining d-separation, exogeneity, collision, and confounding. However, the drawing of the model depends entirely on qualitative use of reason and logic (not statistics or data) (Chen & Pearl, 2015; Elwert, 2013).

instrument is applicable to all cases in the data. If there are cases where the instrument might have a unique causal relationship with the independent variable, so that effects are not *monotonic*, then this is another form of confounding calling for model re-specification.

Although focused on experimental research, a meta-analysis of instrumental variable estimates in political science suggests that researchers routinely fail to offer theoretical arguments that the *IV* is: (1) unrelated to unobserved/omitted causes of Y , (2) has no direct (causal) effect on Y , and (3) that the instrument could plausibly affect all cases (Sovey & Green, 2011)⁹. This neglect has grave implications for the trustworthiness of results.

An Application – Opinion and Policy

I use the example of Breznau (2017) modeling simultaneous feedback between public opinion and social spending to provide a didactical picture of SFMs. I only briefly summarize the theory from the original research, to keep the focus on execution of the SFM. Public opinion and social policy are an example of theoretical simultaneous feedback, because: (1) Opinion and policy are co-determinant occurring at the same moments or overlapping moments in time. Observing public opinion in a one-year unit prevents observation of anything other than simultaneous effects, even if multiple effects take place within a year. (2) The relationship is so complex that a simultaneous model may come closer to reality than something with arbitrary lags (as taken from years of a survey). Policymakers imagine opinion or act on expected future changes in opinion before opinion changes occur, while public opinion responds to policymakers' intentions and discussions before they actually change policy. Moreover, opinion responds to many things at once over many points in time and the responses take different lengths to materialize. The same applies to policymaking. Given all these effects starting, maturing, declining and then stopping over time, I expect that there is a simultaneous effect, or average simultaneous effect underlying all effects.

The instruments I employ are female labor force participation (IV_1) for public opinion (Y_1) and veto points (IV_2) for policy (Y_2). Labor force participation influences policy attitudes. Holding male participation roughly equal (as seen across OECD countries), variation in the distribution of female participation links to changes in aggregate opinion. Women, who are significantly more supportive of social policy than men are, become less supportive when in the labor force, on average. Moreover, the policy 'styles' of different countries show no patterning by female labor force participation suggesting that at least in recent decades it has no effect on social policy in the aggregate (i.e., exogenous from Y_2). Veto points deter-

9 An argument I am guilty of not making in Breznau (2017)!

mines how easy it is to block legislation in the design of the political system (e.g., executive or minority veto, bicameralism or federalism), thus where veto points are higher, policy provisions should be lower. Veto points are part of a larger institutional framework of societies that might influence public opinion; however, previous research suggests that they are independent (i.e., exogenous from Y_1). Moreover, veto points predate the measurement of public opinion by decades if not centuries, further meeting the exclusion restriction (see Breznau, 2017).

The data I use are publically available; public opinion in the *International Social Survey Program* ‘Role of Government’ and ‘Religion’ modules and social policy spending from the *Organization for Economic Co-operation and Development* ‘Social Expenditures Database’ covering 70 country-time points (across 1985-2006). I provide the variances and covariances necessary to estimate the main models. I include means only for didactic purposes (see Appendix 1-Table A1). All variable measurements and countries are in Appendix 1-Table A2, reproducing Breznau (2017, p. 597). Almost all SEM software reads raw data or covariance matrix data (including correlation/variance matrices). Appendix 1-Table A3 provides programming code (some call this “syntax”) for *Mplus*, *Stata* and *R* (*RStudio* running *lavaan*). *Stata* and *R* allow programming the matrix by hand, and *Mplus* reads a .dat file, which is a product of copying the matrix into a text editor and saving it with the file extension .dat¹⁰.

I analyze models of opinion and policy reflecting Figure 3B with four independent X variables (aged population, right-party power, unemployment and GDP) predicting both Y outcomes. Table 1 presents results for M1, a model of free estimation with little theory and no additional model constraints. Column “b” are unstandardized (‘metric’) coefficients, and “ β ” standardized coefficients. The results from *Mplus* here are identical to the other software except rounding error.

The results reveal how much Y_1 and Y_2 cause or explain each other’s variance. The standardized coefficient for Y_2 predicting Y_1 suggests that social policy has a *very large* impact on public opinion (0.715), larger than public opinion has on social policy (0.084). However, according to standard testing the effects are insignificant. The insignificance of the smaller effect is perhaps not surprising but insignificance of the very large effect demonstrates the difficulty in disentangling reciprocal effects statistically. Moreover, the countries are not exactly a sample of a larger population, like with human populations. Cut-offs (e.g., $p < 0.05$) are perhaps arbitrary without a sample population to generalize into. The t-statistic is still useful for gauging the coefficients. Thus, Y_2 impacting Y_1 is more reliable and precise ($t = 0.148 / 0.088 = 1.682$) than vice-versa (at 0.357).

10 A1-Appendix One is at the end of this document. The long-form of all code, data, and supplementary analyses are available in Appendix Two and Three, A2 and A3 at <https://osf.io/gyz6p/>, and .dat files at <https://osf.io/cxzj6/>.

Table 1 Results from M1. Freely Estimated Simultaneous Feedback between Opinion and Policy

Y_1 (public opinion) ON	b	s.e.	β	Fig 3B label
Y_2 (social policy)	0.148	0.088	0.715	b
X_1 (aged)	0.024	0.116	0.052	c₁
X_2 (right)	-0.659	0.656	-0.133	c₂
X_3 (unemp)	-0.070	0.039	-0.264	c₃
X_4 (GDP)	-0.055	0.024	-0.287	c₄
IV_1 (FLP)	-0.073	0.018	-0.540	e
Y_2 (social policy) ON				
Y_1 (public opinion)	0.403	1.129	0.084	a
X_1 (aged)	1.134	0.318	0.507	d₁
X_2 (right)	-4.615	2.560	-0.194	d₂
X_3 (unemp)	0.187	0.140	0.145	d₃
X_4 (GDP)	0.113	0.140	0.124	d₄
IV_2 (veto)	-7.509	2.988	-0.235	f
	<i>variance</i>		<i>std.variance</i>	
e.Y ₁	0.630	0.323	0.654	g
e.Y ₂	13.211	2.242	0.592	h
	<i>covariance</i>		<i>correlation</i>	
(e.Y ₁ ,e.Y ₂)	-1.878	1.213	-0.651	m

Note. b are metric and β are standardized coefficients; 70 country-time point cases from ISSP, OECD and other data sources (see A1-Table A2 or Breznau 2017, M10B); Figure 3B contains only one *X* variable so labels include a subscript to differentiate the four *X* variables in this model.

Scholars should exercise caution when interpreting effects independently. The relationship is a loop, not a single causal arrow. Here this loop accounts for (0.715*0.084=0.06) 6% of the joint distribution of the two *Y* variables (although this percentage also depends on the signs and scaling of the coefficients, see section “Explaining variance”). If correctly specified, social policy is a stronger component of this loop. In fact, the term *field* better describes this relationship because the forces are simultaneous and constant like magnets. The coefficients represent constant forces in this stable field. This contrasts with a cyclical loop where a change in one variable sends effects looping through *Y*₁ and *Y*₂ in a cyclical process. A steady-state force of the loop and a cyclical force running through the loop are different. To say that the levels of *Y*₁ on *Y*₂ are at equilibrium because of their perpetual effects

on each other is different than stating that causal effects between Y_1 on Y_2 unfold in specific, precise periods.

I do not rule out the cyclical version of feedback, but have specific theoretical arguments for a non-cyclical version, one that takes place without yearly-time consideration and is sufficiently complex to warrant SFMs. I might take interest in the cyclical relationship when investigating a specific social policy with specific time periods of voting or policymaking. But this macro-comparative exercise presumes that the sum of all specific instances contains common simultaneous feedback; i.e., not particular to one country-year. The comparative advantage here is the ability to test if the general process formulated in a theory of simultaneous feedback and positive returns can be explained by these data (Breznau, 2017; Pierson, 2000).

Without acknowledging reciprocal causality in some form, scholars might measure a unidirectional effect of Y_1 on Y_2 and then separately estimate unidirectional Y_2 on Y_1 rather than a SFM. Appendix 1-Table A4 reveals results from separate regressions. The striking difference is that in both unidirectional regressions the β -coefficients for Y_1 and Y_2 are close to 0.1. This approach leads researchers to conclude that either public opinion explains or causes social policy (Y_1 causes Y_2) or vice-versa (Y_2 causes Y_1), and in either case that the effect is around magnitude of 0.1 standard deviations. Given a theory of simultaneous or reciprocal causality, both conclusions are false and these models are misspecified¹¹. The theory used in constructing M1, and the non-zero loop effect of 6% are evidence of this misspecification.

Hypothesis Testing – The SEM Perspective

All parameters in M1 are free, showing how causal effects *might* look if I know nothing theoretically about Y_1 and Y_2 feedback. Given a sufficiently detailed theory of simultaneous feedback, a scholar knows something about the feedback. Thus, I test hypotheses derived from this knowledge. This is the structural equation modeler perspective focusing on overidentified models (Bollen, 1989). This perspective aims to test if a hypothetically derived model leads to something not far off from observational data. If the implied covariances of an overidentified model are not significantly different from observed covariances, then the hypothetical model may reflect the real-world data-generating processes. Testing hypotheses means comparing models with different exclusions or constraints to determine which fits the data better. Both model testing and model comparison require overidentified models.

11 For example, Zhu and Lipsmeyer (2015) use ISSP data to show an impact of policy on opinion while Brooks and Manza (2006) use ISSP data to show an impact of opinion on policy without acknowledging reciprocal causality in their models.

Adding more instrumental variables achieves overidentification, as each adds one degree of model freedom. However, instrumental variables are rare and having two here represents the current limits of this research, beyond speculation (Breznau, 2013, p. 132; 136).

Fixing Parameters

Arguments for a reciprocal relationship of Y_1 and Y_2 , are likely to include theory of what this relationship looks like. This is true for opinion and policy feedback (Pier-son, 2000; Soroka & Wlezien, 2010). Thus, I specify hypotheses about the nature of the feedback and fix parameters to reflect this. The methodological advantage is an overidentified model. The theoretical advantages are testing competing hypotheses to construct improved theory.

After reviewing the literature I determine that a thermostatic feedback theory suggests that the standardized coefficient **a** (from Figure 3B) is negative 0.05 and **b** is positive 0.30 (see Breznau, 2017). I fix the parameters to these values in M2. The SEM software analyzes only unstandardized effects, thus it is necessary to derive them by scaling the standard deviation of the standardized variable from one to its observed value¹². Meanwhile an increasing returns theory suggests that both coefficients are positive, possibly around 0.15 as specified in M3. The code is in Appendix 1-Table A5, and Table 2 presents the results.

The other variables' coefficients do not carry much in the way of hypothesis testing (that comes in "Fit testing and diagnostics"); however, they should match theoretical expectations. For example, if the coefficient for aged (X_1) was large and negative, I would become very suspicious that my model is misspecified because it is well-established that more older persons in a society requires far more social spending and usually means greater support of social spending.

A researcher might wish to fix an error term, covariance or mean instead of an effect. M4 has a fixed Y_2 error variance of 0.3, fixed covariance of Y_1 and Y_2 error terms at zero and means of Y_1 and Y_2 at zero. I do not have theoretical arguments for these constraints, they are didactic. Survey data provide the possibility to calculate measurement error for public opinion and I invent the number 0.3 here to represent this possibility. A fixed covariance of zero would be that the model represents a closed system accounting for all possible causal pathways between the variables. This would meet an experimental ideal, where the model explains all things that cause Y_1 , Y_2 and the causal loop between them. But this is highly unlikely in the complex realm of cross-national survey research (see section "Instrumental vari-

12 Standardized effect formula: $\beta = b * \frac{\sigma_x}{\sigma_y}$; metric effect formula: $b = \beta * \frac{\sigma_y}{\sigma_x}$; where β = standardized coefficient, b = metric coefficient, σ_x = standard deviation of the independent variable, and σ_y = standard deviation of the dependent variable.

Table 2 Models of Competing Theories of Opinion-Policy Simultaneous Feedback

variable	M2			M3			M4		
	b	s.e	β	b	s.e	β	b	s.e	β
<i>Y₁</i> (public opinion) ON									
<i>Y₂</i> (social policy)	-0.010	--	-0.048	0.030	--	0.146	0.030	--	0.165
<i>X₁</i> (aged)	0.216	0.038	0.466	0.167	0.037	0.362	0.209	0.027	0.484
<i>X₂</i> (right)	-1.434	0.413	-0.291	-1.240	0.402	-0.252	-1.055	0.331	-0.229
<i>X₃</i> (unemp)	-0.034	0.028	-0.129	-0.044	0.027	-0.165	-0.006	0.018	-0.023
<i>X₄</i> (GDP)	-0.053	0.020	-0.281	-0.053	0.019	-0.282	-0.044	0.016	-0.249
<i>IV₁</i> (FLP)	-0.063	0.015	-0.471	-0.066	0.014	-0.494	-0.045	0.009	-0.358
<i>Y₂</i> (social policy) ON									
<i>Y₁</i> (public opinion)	1.500	--	0.311	0.750	--	0.154	0.750	--	0.137
<i>X₁</i> (aged)	0.901	0.211	0.403	1.062	0.207	0.474	1.175	0.164	0.495
<i>X₂</i> (right)	-3.376	2.264	-0.142	-4.217	2.225	-0.177	-3.929	2.211	-0.156
<i>X₃</i> (unemp)	0.172	0.142	0.134	0.183	0.140	0.142	0.245	0.121	0.180
<i>X₄</i> (GDP)	0.210	0.104	0.229	0.148	0.103	0.160	0.201	0.080	0.206
<i>IV₂</i> (veto)	-8.070	3.107	-0.252	-8.369	2.986	-0.261	-7.183	2.987	-0.212
e. <i>Y₁</i>	0.446	0.075	0.466	0.424	0.072	0.445	0.300	--	0.360
e. <i>Y₂</i>	13.702	2.318	0.613	13.234	2.240	0.589	13.370	2.260	0.532
(e. <i>Y₁</i> ,e. <i>Y₂</i>)	-0.279	0.307	-0.113	-0.472	0.293	-0.199	0.000	--	0.000

Note. *Stata* results shown; *R* (*lavaan*) and *Mplus* identical except rounding error. M4 is not theoretical, has didactic purpose only.

ables”). Nonetheless, I constrain it here for exercise. Means at zero is not important theoretically, it just centers the expected values of *Y₁* and *Y₂*¹³.

Fit Testing and Diagnostics

Tests of fit determine how well a theoretically derived model explains real-world observations or compares with alternative models. There is a small universe of these tests. The art of ruling out alternative theoretical models is crucial to scientific utility (Hayduk et al., 2007; and discussed on the structural equation modeling listserv SEMNET), and primarily comes from investigation of how close the

13 Researchers may have a theory that effects **a** and **b** are equal, but not have any prediction about their size. It is possible to constrain **a** and **b** to equality and let computer estimation decide what size is ideal in all three softwares (see A3-Appendix Three).

model-implied covariances come to the freely observed covariances in the data. The proportion of explained variance (r^2) is often a secondary concern. The term *residual* denotes the differences between model-implied covariances and observed covariances. *Residual* also describes OLS error (in \hat{Y}), thus structural modelers sometimes use *fitted residuals* or *covariance residuals* to adjudicate these concepts (Kline, 2011).

For just-identified models (like M1) the covariance residuals are zero as implied and observed are identical. In overidentified models, larger residuals suggest worse local fit. Scholars rely on *standardized residuals* and *normalized residuals* given that residuals on their own do not have a common metric. Appendix 1-Table A6 provides residuals for M2 and M3. Smaller residuals support M3.

I might worry about the -1.28 normalized residual of IV_2 and Y_1 in M2 (Appendix 1-Table A6). This suggests unexplained covariance remaining between these variables, where none should be present. This might evidence a weak instrument. However, M3 is the preferred model where this residual is slightly lower at -0.964. Given that M3 fits well overall (as shown in Table 3), and that the theory supports the instrument of veto points being exogenous to public opinion, I tentatively defend IV_2 . Yet future research should search for other IV s. What causes policy changes that does not cause opinion changes is a puzzle. Finding strong and valid instruments is a perpetual concern (Antonakis et al., 2010).

The *model chi-square* (χ^2) provides the primary statistic for evaluating global model fit. The χ^2 comes from maximum-likelihood estimation (for a good introduction see Kline, 2011, p. 199). The *exact fit hypothesis* is that implied and observed covariance matrices are identical except for random error. Put into test terms, χ^2 difference should not be significant at $p < 0.05$, otherwise the matrices in comparison are significantly different offering evidence to reject this model. Thus, $p > 0.05$ is a reasonable level to not reject the exact fit hypothesis. If this test passes, it does not guarantee the strength of the IV , but asserts that nothing about the model radically departs from the observed data; i.e., displays reasonable global fit. The exact fit test becomes increasingly likely to fail the larger the sample because it is more likely to pick up very small confounding parameters in the empirical realm. In macro-comparative survey research, having too large of a country sample is unlikely a problem. The *equal fit hypothesis* is that two implied covariance matrices do not differ from one another. If $p < 0.05$ they are significantly different supporting the larger model (with less degrees of freedom). Note that models are only comparable with an equal fit test when they are *nested*; i.e., have all the same basic parameters and observational data.

There are several other global fit diagnostics. Considering all of them is helpful in selecting models, especially when they are not nested (Kline, 2011)¹⁴. Table

14 David Kenny's website provides discussions of model fit <http://davidakenny.net/cm/fit.htm>.

Table 3 Model Fit Statistics and Tests

Statistic	Test	Interpretation	Arguments	M1	M2	M3	M4
χ^2				0	5.456	2.758	18.078
df		Model df	0=just identified	0	2	2	6
P-value	Exact fit	Significance of implied and observed covariance differences	p>0.05 equal covariances	NA	0.065	0.252	0
P-value ^a	Equal fit (cf. M4)		p<0.05 smaller model is worse	NA	0.013	0.004	NA
AIC		lower is better	decrease of 5-10 better	1957.1	1958.6	1955.9	1967.2
BIC		lower is better	decrease of 5-10 better	1990.8	1987.8	1985.1	1991.9
RMSEA		lower is better	good <0.05, bad>0.10	0	0.157	0.074	0.170
P-value	P-close	significance of one-sided test RMSEA is greater than 0.05	p>0.05 rejects	NA	0.097	0.311	0.016
CFI		closer to 1.0 better	good >0.95	1	0.959	0.991	0.855
TLI		closer to 1.0 better	good >0.95, problem >1.0	1	0.731	0.941	0.686
SRMR		standardized difference of implied and observed correlation residuals	0 = identical	0	0.028	0.020	0.048

Note. AIC “Akaike’s Information Criterion”, BIC “Bayesian Information Criterion”, RMSEA “Root Mean Square Error of Approximation”, CFI “Comparative Fit Index”, TLI “Tucker-Lewis Index”, SRMR “Standardized Root Mean Square Residual”

^a Equal fit test uses χ^2 and degrees of freedom statistics as the difference of the current model from M4, then a χ^2 table reveals significance levels, or researchers can use an online calculator or Excel command “=CHISQ.DIST.RT(χ^2 ;df)”

3 contains fit and diagnostics for models M1-M4, offering some preferable targets of these indices. I conclude that M3 is better than M1 because M1 does not have a strong theory to test and AIC and BIC are worse; and better than M2 because all fit indices (AIC, BIC, RMSEA, CFI and TLI) are better. Also, exact fit is less significant (0.252 vs. 0.065) and equal fit more significant (p-value 0.004) than M2 (0.013). It is better than M4, although M4 is just for example.

In addition to residuals, another tool to identify local misfit is *modification indices*. For every parameter in the model, the modification index is the change in χ^2 if that parameter (coefficient or residual covariance) were freely estimated instead of estimated in its current form. The values are zero for parameters already freely estimated and take on positive values for parameters currently fixed (for example the effect of IV_1 on Y_2 in all of the models). Appendix 1-Table A7 lists all non-zero modification indices for M2 and M3. Appendix 1-Table A7 suggests that estimating a free parameter for the regression of Y_2 on IV_1 is a way to improve the model. The normalized residual between Y_2 and IV_1 is -1.28 (see Appendix 1-Table A6) supporting this claim; however, a much larger gain in model fit would result from adding a freely estimated coefficient for Y_1 on IV_2 (4.374 in M2) than for Y_2 on IV_1 (0.745 in M2). This distinction is not evident from looking only at the residuals. Yet, neither of these is possible because the model is not identified with the addition of either parameter (as per the rank condition discussed earlier). Here again are the current limits of this research.

Modification indices are agnostic statistical scores; they do not identify a theoretical problem. Thus, simply freeing parameters in the model might defy, disrupt or debunk the causal model that the researcher carefully constructed using theory. Modification indices are a tool for researchers to use to re-visit their theories and discover what might be missing logically, *before* making any changes to the model. Focusing on M2: In Table A7, the modification indices are identical for the effect of IV_2 on Y_1 and Y_2 on Y_1 , and identical for IV_1 on Y_2 and Y_1 on Y_2 . This demonstrates how endogeneity works in the SFM. There is residual covariance between Y_1 and Y_2 (normalized value of 0.197 in M2) and the fit of the model may suffer as a result, as the modification index of 4.374 suggests. This essentially means there is a statistical relationship (covariance) between Y_1 and Y_2 not explained by the model and if something could account for this unique feedback error, the model would fit better; in this case a better or additional instrument for IV_2 . I did not discuss this in Breznau (2017), but this is a useful finding from this excursus pointing at further research.

Explaining Variance

Sometimes a purpose of explaining variance arises in addition to fit testing. In a SFM, this is a difficult conceptual task. The loop is *the product of both coefficients*

(effects **a** and **b** in Figure 3B and Table 1) running between Y_1 and Y_2 . In M1, the *loop causal effect* of Y_2 on Y_1 is not 0.715, but includes the effect of Y_1 on Y_2 of 0.084 as an indirect effect, and thus $(0.715 \times 0.084) = 0.06$. To calculate this effect as a percentage, take $1/(1 - Y_1 \times Y_2) = 1/(1 - 0.06) = 1.064 =$ the original amount plus 6.4% (Paxton, Hipp, & Marquat-Pyatt, 2011). One cycle through the feedback loop produces about 6.4% of the endogenous variables' covariance¹⁵. To this loop causal effect we may apply a Sobel-like test revealing a significance score (z-value) of 0.131¹⁶. Interpretation is identical to a t-test making this statistic non-significant, which is not surprising given that the coefficients are not significant. Normally, another cycle would recover an additional 6% of 6% of the original covariance and so forth. *In SFMs, there is no perpetual looping effect*. One loop is the theoretically specified 'number of cycles' for the SFM (Hayduk, 2009). The ideal model M3 has a loop causal effect of 2.25% ($= 0.03 \times 0.75$), lower than the 6% found in M1, but offering the best theoretical loop causal effect from this research based on fit diagnostics.

The loop causal effect only offers the amount of unique covariance explained by the loop. The remainder may be of interest to the researcher; however, the amount of explained variance of Y_1 and Y_2 , like their path coefficients, are reciprocally related¹⁷. The error of either Y variable actually contains part error and part non-error coming directly from the other endogenous variable's error and thus violating the definition of error in OLS regression. The non-error part is not a component of the theory underlying the model, but an implication of the feedback loop.

Hayduk (2006) proposes a re-specification of r^2 to resolve this problem called the blocked-error-r-square (be R^2). Perfectly appropriate for SFMs, it equals the percentage of variance explained by the model when excluding the other error term as predictor (i.e., the non-error). The be R^2 in M2 is $(0.517/0.959) = 0.539$ or 53.9% for Y_1 and $(9.887/22.366) = 44.2\%$ for Y_2 , and for M3 the values are 56.1% for Y_1 and 41.7% for Y_2 (see A3-Appendix Three). The results say little about differences between the models; in fact, they point out that modeling two very different theoret-

15 The formula accounts for situations with opposite signed coefficients, or coefficients greater than one. As in any statistical model, all indirect effects should be calculated from unstandardized coefficients, thus the loop causal effect is $(0.148 \times 0.403) = 0.06$. Although the causal effect should be identical regardless of calculation method, always rely on unstandardized ('metric') coefficients.

16 The standard error (SE) of loop causal effect (where the two causal paths **a** and **b** from Figure 3B are subscripted and normal font "b" is a metric coefficient) is: $SE_{ab} = \sqrt{b_a^2 SE_a^2 + b_b^2 SE_b^2}$; the significance test is then $b_a b_b / SE_{ab}$.

17 Although beyond the scope here, it is interesting to think about the direction of this residual covariance. In infinite looping cycles, a negative covariance approaches zero while a positive covariance explodes towards infinity. In the SFM, there is only one cycle, but there is an implied force of direction suggesting that unobserved causes push away from equilibrium (positive) or towards it (negative).

ical perspectives leads to similar explained variances. Given the small sample-size-to-variables-ratio, it is not surprising that these models explain so much variance.

I did not discuss this in Breznau (2017), that simultaneous feedback accounts for just over 2% of the joint distribution of public opinion and social spending. This would be trivial in standard r-square logic, but this is literally the explained variance unique to the loop itself. The feedback loop is like its own independent variable explaining variance in Y_1 and Y_2 . Moreover, this begs the question: what is the loop? It represents the simultaneous impact of public opinion and social policy on one another. This simultaneity occurs in roughly one-year observation windows. Adding more observations should not change this if the loop is stationary at equilibrium. Therefore, disturbances to opinion or policy at best impart a 2% shift in the distribution of opinion and policy. If speaking in terms of majority elections this could make the difference in outcomes. In terms of social spending, this would impart an increase of 60 units (Dollars, Euro, Yen, etc) if a social benefit provides 3,000 units for something (pension, unemployment, etc). These potential outcomes suggest 2% may be non-trivial.

Further Considerations

Estimators

The task of the estimator is to identify what results most closely fit the implied covariance matrix to the observed covariance matrix (Myung, 2003). The most common estimator for this task is maximum likelihood (ML), or one of its many variants. In econometrics instrumental variables estimation often involves two- or three-stage least squares (2SLS or 3SLS) estimators. For SFMs, ML is the least biased estimator because it takes into consideration all information in the system (i.e., both equations) simultaneously. However, misspecification can lead ML to larger bias than 2SLS under some conditions (Paxton et al., 2011). This potential tradeoff suggests that the researcher may gain from running sensitivity checks with 2 or 3SLS to identify misspecification (Kirby & Bollen, 2009), but should not use the results because they are counter to a theory of simultaneity. 2SLS violates the assumption that the errors are correlated (**m** in Figure 3) because it removes the error through instrumental variable stages. However, as noted long ago by economists, any adjustment to one outcome variable or its error term feeds back into the other and estimating the equations separately misses this process (Hausman, 1983, p. 194; Pearl, 2015).

The key is whether unobserved causes (and effects) are randomly distributed with respect to the reciprocally causal relationship of Y_1 and Y_2 . If they are not, then the researcher can have little faith in the estimation of **a** and **b** in Figure 3,

and should reconsider the formal model rather than worrying about estimators. The default in all three software packages and the default for researchers should be ML.

Disequilibrium

If there are meaningful changes in the size or direction of a causal force during the observation period, then SFMs may not be the appropriate tool. Kaplan, Harik and Hotchkiss (2001) demonstrate some risks associated with estimation under disequilibrium. They simulated different systems that experienced a shock before moving back to equilibrium. They took cross-sections out of the data series to estimate SFMs to test the severity of violating the equilibrium assumption. Their findings reveal that both regression coefficients representing the causal effects between endogenous variables (c.f., Y_1 and Y_2 herein) change somewhat dramatically as the system goes from the shock toward its equilibrium point. The error terms follow a similar pattern. The change in size of coefficients is gradual and smooth in the case of systems that move toward equilibrium without major fluctuations; however, when simulating a system with big oscillations the changes to the regression coefficients are sporadic if not chaotic. In either case, the problem is non-ignorable.

A researcher could mistakenly estimate model Figure 2A when in fact the correct model is 2D wherein $Y_{1,t-i}$ shapes $Y_{1,t-1}$ which leads to a new cycle of effects between Y_1 and Y_2 , and then $Y_{1,t-1}$ takes on an entirely new causal effect on $Y_{1,t}$ because of whatever transpired in the first loop (arrows between Y_1 and Y_2) at $t-1$. This means that the model is *cyclically recursive* instead of nonrecursive (Billings & Wroten, 1978). Unfortunately, it is not possible to test for equilibrium, because the data needed for such a test are missing by definition. This leaves a strong burden on the researcher to argue for equilibrium. In the case of macro-comparative survey research, useful arguments may arise based on stable political and cultural systems. For example, the welfare states of Western Europe show a strong degree of stability in their political systems after the 1950s; whereas the Communist states of Eastern Europe broke down and experienced the shock of market transition in the 1990s.

In cross-sectional survey data, there are somewhat random assortments of countries and time-periods available, case-in-point are ISSP data. If the effects and system are truly at equilibrium, then it does not matter what random assortment of country-time-points are in the analysis. All should reveal the same effects. Subdividing the sample, it is possible that the timing of surveys provides a sensitivity test. I demonstrate this by splitting the data into all observations prior to 1998 (Group 1) and all those 1998 and later (Group 2) (see Appendix 1-Table A8 for covariances). I run M2 and M3 separately on the split data. Table A9 reveals that M3 is still preferable to M2 in both groups, and that most effects follow similar patterns between the groups. However, the models do not fit nearly as well as when run on

the pooled data – as seen from a few basic fit indices. Nonetheless, the χ^2 p-value from the exact fit tests passes and it appears reasonable that effects are stable over time, for all non-missing years. The very small sample sizes are likely to blame for the troubling other indicators. I compare implied covariance matrices for M3 in Appendix 1-Table A10. Here the main test variables in the model (involving IV_1 , IV_2 , Y_1 and Y_2) carry similar implied covariances across the two groups. A potential problem is X_4 (unemployment), which switches signs for some of the covariances between the groups. This is evidence that further consideration should be given to this variable in future research to see if it is disrupting the stability of the system. Also, *maybe* there was a slightly different size of effects in Group 1 given the model fits the Group 2 data better; although, much more work is necessary here. This sensitivity analysis does not guarantee stability, and although this procedure is not an established method, it follows the art of structural equation modeling to pay detailed attention to model diagnostics.

Other Concerns

Missing values. Strictly speaking missing values should be dealt with in the estimation of the model as opposed to imputing them separately as if they were observed values. The reason for this is that missing values are subject to special measurement error and ignoring this can produce misleading results. However, contextual-level data are not observations in the strict sense of the word. Values for gross domestic product or level of democracy for example stem from complex calculations whose inputs are not necessarily identical across societies. Researchers at organizations such as the OECD take painstaking efforts to make these values as identical as possible. These values do not represent objective qualities of societies in the way that observed variables such as age or height represent objective features of individuals. Contextual variables are instead more abstract. If they are missing it is best to take the nearest available year. The SFM is not suited for imputing values because of endogeneity.

Aggregation and Comparison. Survey data come from micro-level observations, but macro-comparative researchers aggregate them in some way. Researchers should identify population averages, and then use weights and appropriate measurement models, perhaps performing aggregation in several ways as sensitivity analyses. Monte Carlo simulations suggest that idiosyncratic research practices related to weighting and measurement easily impact results in small-N studies (Breznau, 2016). Furthermore, in order to meaningfully use comparative survey data, all questions need the same cognitive meaning in each socio-cultural context (Davidov et al., 2014). Researchers should establish measurement invariance before using survey data, and correct for measurement error using a measurement model

and predicted latent scores that account for differential item functioning when there are three or more scale variables. In this example, previous research suggests measurement invariance of the two ISSP questions (Andreß & Heien, 2001)¹⁸. Given that there are only two items, the loadings are equal. Thus, a predicted 'factor' is identical in variance with simply taking their mean as I did here.

Estimation without instruments. Several authors suggest estimating IV models without observed instrumental variables. Theoretically speaking this violates the exclusion restriction. These methods include estimating a latent or model-implied instrumental variable, or finding a subgroup of the total sample where a researcher can identify a causal instrument (Bollen, Kolenikov, & Bauldry, 2014; Ebbes et al., 2005; Heckman, Urzua, & Vytlačil, 2006; Heckman & Vytlačil, 1999). Suffice to say it is possible but not recommended.

Nonlinear models. If the endogenous variables are non-linear, SFMs are still possible using alternative regression estimation techniques. Simply resorting to linear probability models may introduce new forms of bias (Finch & French, 2015; Terza, Bradford, & Dismuke, 2008)

Conclusion

This excursus shows that data limitations of macro-comparative research are not always a burden. With a theory of sub-yearly causal timing, scholars need not automatically reject cross-sectional survey data as a source for investigating their hypotheses. There are many theoretical forms of reciprocal causality for this. The simultaneous feedback model is only one form. Awareness of this method is not a sufficient condition to use it. Every step in the process of modeling simultaneous feedback must have theoretical argumentation behind it. Theory is a necessary condition for employing a simultaneous feedback model. Without a theory to specify the model, there is no identification of the reciprocal effects and probably no identification of the model. Instrumental variables do not appear through random chance or out of thin air. Perhaps those normally running a bunch of correlations or regressions and then trying to explain the results may learn something from simultaneous feedback modeling, because theory is not 'optional' (Kalter & Kroneberg, 2014).

The impetus for bringing light to this method is the fact that so many macro-comparative phenomena in survey research appear to have reciprocal causality, and the forms of causality are highly complex and unfold in imprecise moments in time. There are well established methods, for example cross-lagged, fixed-effects/random-slope, error correction and vector autoregressive models for fitting longitu-

18 Others find similar questions to be measurement equivalent in the ESS (Roosma, van Oorschot, & Gelissen, 2014)

dinal models. Given the correct research design it is possible to integrate simultaneous feedback in a longitudinal model (Geweke, 1982) like an extension of Figure 2D. Whether or not simultaneous feedback can capture both lagged and instantaneous processes is a theoretical consideration, one limited by available data. The loop causal effect from a SFM may then impact other outcomes (Hayduk, 1987). The loop itself acts as an independent 'variable' or a causal force, a consideration that researchers hopefully take away from this excursus.

There are limitations. Although data derive from individual-level sources, I am not aware of the possibility to model a SFM using multi-level techniques nor individual-level measurement models. Ideally, a measurement model is integrated into a path model for a fully parsimonious structural equation model. This would have a single variable for each survey item and their relationship with the latent scale (here public opinion), and it would have two levels of data analysis. Lacking degrees of freedom prevents the former, and a peculiarity of the SFM prevents the latter. The loop only exists at the aggregate level because there is no individual-level variance in social policy. Moreover, public opinion is by definition a group-level phenomenon, meaning strictly macro-level.

Theories germane to simultaneous feedback come in two broad types and both are debatable, so that researchers should use caution. The first type is where forces act upon each other simultaneously in the real world. The possibility of this is a philosophical argument. Some argue that by definition there are actions and reactions in the world, or that all things are reactions to other things. Meanwhile others argue that it is the interaction of objects and actions at the same point in time that constitute causal effects (Mulaik, 2009). Although this paper takes no philosophical position, researchers working with SFMs are by definition stepping on philosophical ground and tapping into debates that stretch throughout the history of social thought. Thus, awareness of these arguments should help researchers defend themselves against epistemological attacks. The second type suggests that simultaneous causality exists without theoretically simultaneous forces, but can be inferred because the window of observation – usually something around a year in surveys – contains enough bi-directional causal forces between two phenomena that it is logical to treat them as simultaneously causal. This means that even though all these effects may run in different directions and have different sizes, that there is a sum or total effect in their causal loop force that is of theoretical and empirical interest.

Although simultaneity across many countries is an interesting comparative perspective to take and test, researchers more often think of comparative research as looking for differences. As my sensitivity analysis in A3-Appendix Three shows, I can compare two different groups in the data, analogous to a moderation analysis. There are theories that opinion and policy will have different sized effects depending on the institutional context (Wlezien & Soroka, 2012), and this presents an

exciting avenue for future implementation of simultaneous feedback in macro-comparative survey data in general and specifically in the opinion-policy case.

References

- Andreß, H.-J., & Heien, T. (2001). Four Worlds of Welfare State Attitudes? A Comparison of Germany, Norway and the United States. *European Sociological Review*, 17(4), 337–356. <http://doi.org/10.1093/esr/17.4.337>
- Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, 91(434), 444–455. <http://doi.org/10.2307/2291629>
- Angrist, J. D., & Krueger, A. B. (2001). Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives*, 15(4), 69–85. <http://doi.org/10.3386/w8456>
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On Making Causal Claims: A Review and Recommendations. *The Leadership Quarterly*, 21(6), 1086–1120. <http://doi.org/10.1016/j.leaqua.2010.10.010>
- Bartels, L. M. (1991). Instrumental and “Quasi-Instrumental” Variables. *American Journal of Political Science*, 35(3), 777–800. <http://doi.org/10.2307/2111566>
- Bascle, G. (2008). Controlling for endogeneity with instrumental variables in strategic management research. *Strategic Organization*, 6(3), 285–327. <http://doi.org/10.1177/1476127008094339>
- Billings, R. S., & Wroten, S. P. (1978). Use of path analysis in industrial/organizational psychology: Criticisms and suggestions. *Journal of Applied Psychology*, 63(6), 677–688. <http://doi.org/10.1037/0021-9010.63.6.677>
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York, NY: John Wiley & Sons.
- Bollen, K. A. (2012). Instrumental Variables in Sociology and the Social Sciences. *Annual Review of Sociology*, 38(1), 37–72. <http://doi.org/10.1146/annurev-soc-081309-150141>
- Bollen, K. A., Kolenikov, S., & Bauldry, S. (2014). Model-Implied Instrumental Variable—Generalized Method of Moments (MIIV-GMM) Estimators for Latent Variable Models. *Psychometrika*, 79(1), 20–50. <http://doi.org/10.1007/s11336-013-9335-3>
- Brehm, J., & Rahn, W. (1997). Individual-Level Evidence for the Causes and Consequences of Social Capital. *American Journal of Political Science*, 41(3), 999–1023. <http://doi.org/10.2307/2111684>
- Breznau, N. (2013). *Public Opinion and Social Policy*. Bremen, Germany: University of Bremen Library. <https://elib.suub.uni-bremen.de/peid=D00103291>
- Breznau, N. (2016). Secondary Observer Effects: Idiosyncratic Errors in Small-N Secondary Data Analysis. *International Journal of Social Research Methodology*, 19(3), 301–318. <http://doi.org/10.1080/13645579.2014.1001221>
- Breznau, Nate. 2017. Positive Returns and Equilibrium: Simultaneous Feedback Between Public Opinion and Social Policy. *Policy Studies Journal* 45(4), 583–612. <http://doi.org/10.1111/psj.12171>
- Brooks, C., & Manza, J. (2006). Social Policy Responsiveness in Developed Democracies. *American Sociological Review*, 71(3), 474–494.

- <http://doi.org/10.1177/000312240607100306>
- Chen, B., & Pearl, J. (2015). *Exogeneity and Robustness*. Technical Report R-449
http://ftp.cs.ucla.edu/pub/stat_ser/r449.pdf.
- Chong, A., & Gradstein, M. (2007). Inequality and Institutions. *Review of Economics and Statistics*, 89(3), 454–465. <http://doi.org/10.1162/rest.89.3.454>
- Claibourn, M. P., & Martin, P. S. (2000). Trusting and Joining? An Empirical Test of the Reciprocal Nature of Social Capital. *Political Behavior*, 22(4), 267–291.
<http://doi.org/10.1023/A:1010688913148>
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement Equivalence in Cross-National Research. *Annual Review of Sociology*, 40(1), 55–75.
<http://doi.org/10.1146/annurev-soc-071913-043137>
- Ebbes, P., Wedel, M., Böckenholt, U., & Steerneman, T. (2005). Solving and Testing for Regressor-Error (in)Dependence When no Instrumental Variables are Available: With New Evidence for the Effect of Education on Income. *Quantitative Marketing and Economics*, 3(4), 365–392. <http://doi.org/10.1007/s11129-005-1177-6>
- Elwert, F. (2013). Graphical Causal Models. In S. L. Morgan (Ed.), *Handbook of Causal Analysis for Social Research* (pp. 245–272). Dordrecht, The Netherlands: Springer Science & Business Media. http://doi.org/10.1007/978-94-007-6094-3_13
- Finch, W. H., & French, B. F. (2015). Modeling of Nonrecursive Structural Equation Models With Categorical Indicators. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(3), 416–428. <http://doi.org/10.1080/10705511.2014.937380>
- Finkel, S. E. (1995). *Causal Analysis with Panel Data* (ebook / Qu). Thousand Oaks: Sage.
- Geweke, J. (1982). Measurement of Linear Dependence and Feedback between Multiple Time Series. *Journal of the American Statistical Association*, 77(378), 304–313.
<http://doi.org/10.1080/01621459.1982.10477803>
- Hahn, J., & Hausman, J. (2002). A New Specification Test for the Validity of Instrumental Variables. *Econometrica*, 70(1), 163–189. <http://doi.org/10.1111/1468-0262.00272>
- Hausman, J. A. (1983). Chapter 7 Specification and estimation of Simultaneous Equation Models. In B. T.-H. of Econometrics (Ed.), (Vol. Volume 1, pp. 391–448). Elsevier.
[http://doi.org/http://dx.doi.org/10.1016/S1573-4412\(83\)01011-9](http://doi.org/http://dx.doi.org/10.1016/S1573-4412(83)01011-9)
- Hayduk, L. A. (1987). *Structural Equation Modeling with LISREL: Essentials and Advances*. Baltimore: Johns Hopkins University Press.
- Hayduk, L. A. (2006). Blocked-Error-R2: A Conceptually Improved Definition of the Proportion of Explained Variance in Models Containing Loops or Correlated Residuals. *Quality & Quantity*, 40, 629–649. <http://doi.org/10.1007/s11135-005-1095-4>
- Hayduk, L. A. (2009). Finite Feedback Cycling in Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(4), 658–675.
<http://doi.org/10.1080/10705510903206030>
- Hayduk, L. A., Cummings, G., Boadu, K., Pazderka-Robinson, H., & Boulianne, S. (2007). Testing! testing! One, Two, Three – Testing the Theory in Structural Equation Models! *Personality and Individual Differences*, 42(5), 841–850.
<http://doi.org/http://dx.doi.org/10.1016/j.paid.2006.10.001>
- Heckman, J. J. (2000). Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective. *The Quarterly Journal of Economics*, 115(1), 45–97.
<http://doi.org/doi.org/10.1162/003355300554674>

- Heckman, J. J., Urzua, S., & Vytlačil, E. (2006). Understanding Instrumental Variables in Models with Essential Heterogeneity. *Review of Economics and Statistics*, 88(3), 389–432. <http://doi.org/10.1162/rest.88.3.389>
- Heckman, J. J., & Vytlačil, E. J. (1999). Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects. *Proceedings of the National Academy of Sciences*, 96(8), 4730–4734. <http://doi.org/10.1073/pnas.96.8.4730>
- Kalter, F., & Kroneberg, C. (2014). Between Mechanism Talk And Mechanism Cult: New Emphases in Explanatory Sociology And Empirical Research. *KZfSS Kölner Zeitschrift Für Soziologie Und Sozialpsychologie*, 66(1), 91–115. <http://doi.org/10.1007/s11577-014-0272-7>
- Kaplan, D., Harik, P., & Hotchkiss, L. (2001). Cross-Sectional Estimation of Dynamic Structural Equation Models in Disequilibrium. In R. Cudeck, K. G. Jöreskog, & D. Sörbom (Eds.), *Structural Equation Modeling: Present and Future: a Festschrift in honor of Karl Jöreskog* (p. 315). Lincolnwood, IL: Scientific Software International.
- Kirby, J. B., & Bollen, K. A. (2009). Using Instrumental Variable Tests to Evaluate Model Specification in Latent Variable Structural Equation Models. *Sociological Methodology*, 39(1), 327–355. <http://doi.org/10.1111/j.1467-9531.2009.01217.x>
- Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling* (Third). New York, NY: Guilford Press.
- Liska, A. E., & Reed, M. D. (1985). Ties to Conventional Institutions and Delinquency: Estimating Reciprocal Effects. *American Sociological Review*, 50(4), 547–560. <http://doi.org/10.2307/2095438>
- MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the Mediation, Confounding and Suppression Effect. *Prevention Science*, 1(4), 173–181. <http://doi.org/10.1023/A:1026595011371>
- Marini, M. M. (1984). Women's Educational Attainment and the Timing of Entry into Parenthood. *American Sociological Review*, 49(4), 491–511. <http://doi.org/10.2307/2095464>
- Markus, K. A. (2010). Structural Equations and Causal Explanations: Some Challenges for Causal SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(4), 654–676. <http://doi.org/10.1080/10705511.2010.510068>
- Mulaik, S. A. (2009). *Linear Causal Modeling with Structural Equations*. Boca Raton: CRC Press, Taylor & Francis.
- Mulatu, M. S., & Schooler, C. (2002). Causal Connections between Socio-Economic Status and Health: Reciprocal Effects and Mediating Mechanisms. *Journal of Health and Social Behavior*, 43(1), 22–41.
- Myung, I. J. (2003). Tutorial on Maximum Likelihood Estimation. *Journal of Mathematical Psychology*, 47(1), 90–100. [http://doi.org/http://dx.doi.org/10.1016/S0022-2496\(02\)00028-7](http://doi.org/http://dx.doi.org/10.1016/S0022-2496(02)00028-7)
- OECD. (2012). *Social Expenditures Database*. accessed 01.10.2012: Organization for Economic Cooperation and Development.
- Owens, T. J. (1994). Two Dimensions of Self-Esteem - Reciprocal Effects of Positive Self-Worth and Self-Deprecation on Adolescent Problems. *American Sociological Review*, 59(310), 391–407.
- Page, B. I., & Jones, C. C. (1979). Reciprocal Effects of Policy Preferences, Party Loyalties and the Vote. *The American Political Science Review*, 73(4), 1071–1089. <http://doi.org/10.2307/1953990>

- Paxton, P., Hipp, J. R., & Marquatt-Pyatt, S. (2011). *Nonrecursive Models: Endogeneity, Reciprocal Relationships and Feedback Loops*. London: Sage Publications, Ltd.
- Pearl, J. (2010). The Foundations of Causal Inference. *Sociological Methodology*, 40(1), 75–149. <http://doi.org/10.1111/j.1467-9531.2010.01228.x>
- Pearl, J. (2015). Trygve Haavelmo and the Emergence of Causal Calculus. *Economic Theory*, 31, 152–179. <http://doi.org/doi:10.1017/S0266466614000231>
- Pedhazur, E. J. (1997). *Multiple Regression in Behavioral Research: Explanation and Prediction*. New York: Thomson Learning.
- Pierson, P. (2000). Increasing Returns, Path Dependence, and the Study of Politics. *The American Political Science Review*, 94(2), 251–267. <http://doi.org/10.2307/2586011>
- Rigdon, E. E. (1994). Calculating Degrees of Freedom for a Structural Equation Model. *Structural Equation Modeling: A Multidisciplinary Journal*, 1(3), 274–278. <http://doi.org/10.1080/10705519409539979>
- Rigdon, E. E. (1995). A Necessary and Sufficient Identification Rule for Structural Models Estimated in Practice. *Multivariate Behavioral Research*, 30(3), 359–383. http://doi.org/10.1207/s15327906mbr3003_4
- Roosma, F., van Oorschot, W., & Gelissen, J. (2014). The Preferred Role and Perceived Performance of the Welfare State: European Welfare Attitudes from a Multidimensional Perspective. *Social Science Research*, 44C, 200–210. <http://doi.org/10.1016/j.ssresearch.2013.12.005>
- Sargan, J. D. (1958). The Estimation of Economic Relationships using Instrumental Variables. *Econometrica*, 26(3), 393–415. <http://doi.org/10.2307/1907619>
- Schaubroeck, J. (1990). Investigating Reciprocal Causation in Organizational Behavior Research. *Journal of Organizational Behavior*, 11(1), 17–28. <http://doi.org/10.1002/job.4030110105>
- Siegel, S. J., & Alloy, L. B. (1990). Interpersonal perceptions and consequences of depressive-significant other relationships: A naturalistic study of college roommates. *Journal of Abnormal Psychology*, 99(4), 361–373. <http://doi.org/10.1037/0021-843X.99.4.361>
- Sobel, M. E. (1990). Effect Analysis and Causation in Linear Structural Equation Models. *Psychometrika*, 55(3), 495–515. <http://doi.org/10.1007/bf02294763>
- Sobel, M. E. (1996). An Introduction to Causal Inference. *Sociological Methods & Research*, 24(3), 353–379. <http://doi.org/10.1177/0049124196024003004>
- Soroka, S. N., & Wlezien, C. (2010). *Degrees of Democracy: Politics, Public Opinion and Policy*. Cambridge: Cambridge University Press.
- Sovey, A. J., & Green, D. P. (2011). Instrumental Variables Estimation in Political Science: A Readers' Guide. *American Journal of Political Science*, 55(1), 188–200. <http://doi.org/10.1111/j.1540-5907.2010.00477.x>
- Svensson, R., Dahlberg, S., Kumlin, S., & Rothstein, B. (2010). *The QoG Social Policy Dataset*. University of Gothenburg: The Quality of Government Institute.
- Terza, J. V., Bradford, W. D., & Dismuke, C. E. (2008). The Use of Linear Instrumental Variables Methods in Health Services Research and Health Economics: A Cautionary Note. *Health Services Research*, 43(3), 1102–1120. <http://doi.org/10.1111/j.1475-6773.2007.00807.x>
- Thornton, A., Axinn, W. G., & Hill, D. H. (1992). Reciprocal Effects of Religiosity, Cohabitation, and Marriage. *American Journal of Sociology*, 98(36891), 628–651. <http://doi.org/10.1086/230051>

- Wlezien, C., & Soroka, S. N. (2012). Political Institutions and the Opinion–Policy Link. *West European Politics*, 35(6), 1407–1432. <http://doi.org/10.1080/01402382.2012.713752>
- Wong, C.-S., & Law, K. S. (1999). Testing Reciprocal Relations by Nonrecursive Structural Equation Models Using Cross-Sectional Data. *Organizational Research Methods*, 2(1), 69–87. <http://doi.org/10.1177/109442819921005>
- Wright, S. (1920). The Relative Importance of Heredity and Environment in Determining the Piebald Pattern of Guinea-Pigs. *Proceedings of the National Academy of Sciences*, 6(6), 320–332. <http://doi.org/10.1073/pnas.6.6.320>
- Xie, F., & Levinson, D. (2010). How Streetcars Shaped Suburbanization: A Granger Causality Analysis of Land Use and Transit in the Twin Cities. *Journal of Economic Geography*, 10(3), 453–470. <http://doi.org/10.1093/jeg/lbp031>
- Zhu, L., & Lipsmeyer, C. S. (2015). Policy Feedback and Economic Risk: The Influence of Privatization on Social Policy Preferences. *Journal of European Public Policy*, 22(10), 1489–1511. <http://doi.org/10.1080/13501763.2015.1031159>

Appendix 1: Additional Tables

Table A1 Public Opinion and Social Policy Covariance Structure Data^a

		Y_1	Y_2	X_1	X_2	X_3	X_4	IV_1	IV_2
Means		0.085	21.370	14.830	0.150	7.300	25.600	52.000	0.348
Variance		0.976	22.658	4.537	0.040	13.764	26.936	53.729	0.022
Standard Deviations		0.988	4.760	2.130	0.200	3.710	5.190	7.330	0.149
<i>variable</i>	<i>label</i>	Y_1	Y_2	X_1	X_2	X_3	X_4	IV_1	IV_2
correlations	Public Opinion	Y_1	1.000						
	Social Spending	Y_2	0.348	1.000					
	Aged	X_1	0.413	0.532	1.000				
	Right	X_2	-0.141	-0.193	0.052	1.000			
	Unemp.	X_3	0.294	0.128	0.017	0.004	1.000		
	GDP	X_4	-0.405	0.041	0.082	-0.140	-0.525	1.000	
	FLP	IV_1	-0.527	0.003	-0.030	-0.164	-0.585	0.572	1.000
	Veto	IV_2	-0.068	-0.199	0.053	-0.013	-0.064	0.175	-0.191
<i>variable</i>	<i>label</i>	Y_1	Y_2	X_1	X_2	X_3	X_4	IV_2	IV_1
covariances	Public Opinion	Y_1	0.976						
	Social Spending	Y_2	1.638	22.658					
	Aged	X_1	0.869	5.397	4.537				
	Right	X_2	-0.028	-0.184	0.022	0.040			
	Unemp.	X_3	1.077	2.253	0.003	0.130	13.764		
	GDP	X_4	-2.076	1.015	-0.145	0.902	-10.117	26.936	
	FLP	IV_1	-3.818	0.098	-0.241	-0.470	-15.895	21.772	53.729
	Veto	IV_2	-0.010	-0.141	0.000	0.017	-0.035	0.136	-0.209

^a Taken from Breznau (2017).

Table A2 Variable Names and Definitions^a

Name	Type	Measurement	Source
Public Opinion	Endogenous Dependent Variable	Two-item scale from respondents level of agreement with the responsibility of government to provide jobs and reduce income differences.	ISSP Role of Government (I,II,III,&IV) and Religion (I&II) modules ^b
Social Spending	Endogenous dependent variable measuring Social Policy	The amount of spending on social policy provisions, mostly pensions, employment, unemployment, and health care expressed as a percentage of GDP in the same year.	OECD (2012); also known as “SOCX”
Aged	Independent variable	Percent of the population over age 64.	OECD Social Indicators Data
Right	Independent variable	Percent of national government seats held by right parties.	Svennson et al. (2012); Quality of Government Data
Unemp.	Independent variable	Percent of the labor force that is unemployed.	OECD Social Indicators Data
GDP	Independent variable	Gross Domestic Product at PPP.	OECD Social Indicators Data
Female LFP	Instrument for Public Opinion	Percent of the total female population in the labor force.	OECD Social Indicators Data
Veto Points	Instrument for both Social Policy variables	A scale of institutional measures for the amount of chances a policy has to be vetoed. Based on the work of Lijphart (1999).	Svennson et al. (2012); Quality of Government Data

a This Table is copied from Table 1 in Breznau (2017). See original article for full citations. All variables are measured simultaneously at the current year of the endogenous variables.

b Country-time points are: Australia ('86,'90,'93,'97,'98,'07), Austria ('86,'93,'98), Canada ('96,'00,'06), Denmark ('98,'08), Finland ('06), France ('97,'98,'06), Germany ('86,'90,'91,'96,'98,'06), Ireland ('91,'96,'98,'06), Italy ('86,'90,'96,'99), Japan ('96,'98,'06), the Netherlands ('91,'98,'06), New Zealand ('91,'97,'98,'06), Norway ('90,'91,'96,'98,'06), Portugal ('99,'06), Spain ('96,'98,'07), Sweden ('96,'98,'06), Switzerland ('98,'99,'07), Great Britain ('86,'90,'91,'96,'98,'06), the United States ('86,'90,'91,'96,'98,'06).

Table A3 Programming Code for Simultaneous Feedback Models

<i>Mplus</i>	<i>R, lavaan</i>	<i>Stata</i>
Data: FILE IS /data/pospC.dat; !The file pospC.dat must be specified above TYPE IS CORRELATION STDEVIAIONS; !also accepts 'MEANS' and 'COVARIANCE' NOBSERVATIONS ARE 70; Variable: NAMES ARE y1 y2 x1 x2 x3 x4 iv1 iv2;	<pre>library(lavaan) #means in same order as cov matrix m.mean <- ' 0.085 21.37 14.83 0.15 7.3 25.6 52 0.348' #line breaks can go anywhere in the list cov <- '-0.9761 1.6366 22.6576 0.8691 5.3938 4.5369 -0.0279 -0.1837 0.0222 0.04 1.0777 2.2604 0.1343 0.003 13.7641 -2.0767 1.0129 0.9065 -0.1453 -10.1088 26.9361 -3.8166 0.1047 -0.4684 -0.2404 -15.9087 21.7604 53.7289 -0.01 -0.1411 0.0168 -0.0004 -0.0354 0.1353 -0.2086 0.0222' posp.cov <- getCov(cov, names = c('y1', 'y2', 'x1', 'x2', 'x3', 'x4', 'iv1', 'iv2'))</pre>	<pre>clear all ssd init y1 y2 x1 x2 x3 x4 iv1 iv2 ssd set obs 70 ssd set means 0.085 21.37 14.83 0.15 7.3 25.6 52 0.348 ssd set sd 0.988 4.76 2.13 2 3.71 5.19 7.33 0.149 ssd set cor 1 0.348 1 0.413 0.532 1 -0.141 -0.193 0.052 1 0.294 0.128 0.017 0.004 1 -0.405 0.041 0.082 -0.140 -0.525 1 -0.527 0.003 -0.030 -0.164 -0.585 0.572 1 -0.068 -0.199 0.053 -0.013 -0.064 0.175 -0.191 1 *use „ssd set cov“ for covariances</pre>
Analysis: TYPE = GENERAL; !for raw data add the following two lines !MODEL = NOMEANSTRUCTURE; !INFORMATION = EXPECTED;	<pre>m1.model <- ' #Regressions y1 ~ y2 + x1 + x2 + x3 + x4 + iv1 y2 ~ y1 + x1 + x2 + x3 + x4 + iv2 #Correlated Residuals y1 ~~ y2'</pre>	<pre>sem (y2 x1 x2 x3 x4 iv1 -> y1) (y1 x1 x2 x3 x4 iv2 -> y2), cov(e.y1*e.y2) noneans standardized *remove „standardized“ for metric estimates</pre>
Model: y1 ON y2 x1 x2 x3 x4 iv1; y2 ON y1 x1 x2 x3 x4 iv2; y1 WITH y2;	<pre>#this saves the results as an object named „fit“ fit <- sem(m1.model, sample.cov = posp.cov, sample.nobs = 70, meanstructure = FALSE) #lavaan command to display results summary(fit) standardizedSolution(fit)</pre>	
Output: STDYX; !include standardized estimates		

Note. Programmed using *Mplus* 7, *R* (lavaan) 0.5-22 and *Stata* 14; “!” in *Mplus*, “#” in *R*, and “*” in *Stata* are comments; *R* (lavaan) only reads in covariance data (as of 25.03.2017), thus four decimal places used to make estimates as close as possible to the *Mplus* and *Stata* correlation and standard deviation data; variables labeled y1 (social spending), y2 (public opinion), x1 (aged), x2 (right), x3 (gdp), x4 (unemp), iv1 (female labor force participation), and iv2 (veto points).

Table A4 Results from Separate Unidirectional Regressions

Y_1 (public opinion) ON	b	s.e.	β
Y_2 (social policy)	0.029	0.020	0.141
X_1 (aged)	1.789	0.475	0.362
X_2 (right)	-0.117	0.039	-0.252
X_3 (GDP)	-0.044	0.028	-0.165
X_4 (unemp)	-0.053	0.019	-0.280
IV_1 (FLP)	-0.067	0.015	-0.495
$\text{var}(e.Y_1)$	0.423	0.072	0.441
Y_2 (social policy) ON			
Y_1 (public opinion)	0.559	0.581	0.116
X_1 (aged)	11.721	2.573	0.492
X_2 (right)	-0.417	0.217	-0.187
X_3 (GDP)	0.184	0.140	0.144
X_4 (unemp)	0.127	0.114	0.138
IV_2 (veto)	-7.494	2.986	-0.235
$\text{var}(e.Y_2)$	13.197	2.231	0.591

Table A5 Code for Fixing Parameters in the Opinion-Policy Feedback Example (see Tables 2 and 3)

<i>Mplus</i>	<i>R, lavaan</i>	<i>Stata</i>
* * * The upper part of the code for each software is identical to Table 2. * * *		
Model Two (M2)	<pre>Analysis: TYPE = GENERAL; Model: y1 ON y2@-0.01; y1 ON x1 x2 x3 x4 iv1; y2 ON y1@1.5; y2 ON x1 x2 x3 x4 iv2; y1 WITH y2;</pre>	<pre>sem (y2@-0.01 x1 x2 x3 x4 iv1 -> y1) (y1@1.5 x1 x2 x3 x4 iv2 -> y2), cov(e.y1*e.y2) nomeans standardized</pre>
Model Three (M3)	<pre>m2.model <- , #Regressions y1 ~ -0.01*y2 + x1 + x2 + x3 + x4 + iv1 y2 ~ 1.5*y1 + x1 + x2 + x3 + x4 + iv2 #Correlated Residuals y1 ~~ y2' fit2 <- sem(m2.model, sample.cov = posp.cov, sample.nobs = 70, meanstructure = FALSE) m3.model <- , #Regressions y1 ~ 0.03*y2 + x1 + x2 + x3 + x4 + iv1 y2 ~ 0.75*y1 + x1 + x2 + x3 + x4 + iv2 #Correlated Residuals y1 ~~ y2' fit3 <- sem(m3.model, sample.cov = posp.cov, sample.nobs = 70, meanstructure = FALSE)</pre>	<pre>sem (y2@0.03 x1 x2 x3 x4 iv1 -> y1) (y1@0.75 x1 x2 x3 x4 iv2 -> y2), cov(e.y1*e.y2) nomeans standardized</pre>

Mplus	R, lavaan	Stata
<pre>!Change command to include data with means FILE IS /data/pospCM.dat; TYPE IS CORRELATION MEANS STDDEV- VIATIONS; Model: y1 (e1); !labels the y1 error term [y1@0]; !fix mean/intercept to 0 [y2@0]; y1 ON y2@0.03; y1 ON x1 x2 x3 x4 iv1; y2 ON y1@0.75; y2 ON x1 x2 x3 x4 iv2; !remove correlated y1 WITH y2 error, defaults to 0 Model Constraint: e1 = 0.3; !This fixes the e1 variance to 0.3</pre>	<pre>m4.model <- , #Regressions y1 ~ 0.03*y2 + x1 + x2 + x3 + x4 + iv1 y2 ~ 0.75*y1 + x1 + x2 + x3 + x4 + iv2 #remove correlated errors y1 ~~ y2, defaults to 0 fit4 <- sem(m4.model, sample.cov = posp.cov, sample.nobs = 70, meanstructure = FALSE)</pre>	<pre>*variances and means/intercepts fixed as options sem (y2@0.03 x1 x2 x3 x4 iv1 _cons@0 -> y1) (y1@0.75 x1 x2 x3 x4 iv2 _cons@0 -> var(e.y1@0.3) standardized *Remove correlated e.y1 e.y2 errors, defaults</pre>
<pre>Model Fit Output: STDYX !include standardized results SAMPSTAT !observed covariances RESIDUAL !covariance residuals MODINDICES(0); !modification indices, min 0</pre>	<pre>#“object” must be replaced with model name summary(object) standardizedSolution(object) residuals(object) #raw covariance residuals residuals(object, type = „standardized“) residuals(object, type = „normalized“) modindices(object) fitMeasures(object)</pre>	<pre>*remove „standardized“ option to get metric results estat residuals, norm standardized estat mindices, min(0) estat gof, stats(all)</pre>

Note. M2 and M3 correspond to M11B and M12B in Breznau (2017), and M4 is for didactic purposes on fixing parameters other than coefficients.

M2 (Model Two)										M3 (Model Three)								
Normalized Residuals																		
var	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄	IV ₁	IV ₂		label	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄	IV ₁	IV ₂
Y ₁	0.022									Y ₁	0.055							
Y ₂	0.115	-0.008								Y ₂	0.197	-0.039						
X ₁	0.000	0.000	0.000							X ₁	0.000	0.000	0.000					
X ₂	0.000	0.000	0.000	0.000						X ₂	0.000	0.000	0.000	0.000				
X ₃	0.000	0.000	0.000	0.000	0.000					X ₃	0.000	0.000	0.000	0.000	0.000			
X ₄	0.000	0.000	0.000	0.000	0.000	0.000				X ₄	0.000	0.000	0.000	0.000	0.000	0.000		
IV ₁	-0.059	0.442	0.000	0.000	0.000	0.000	0.000			IV ₁	-0.002	0.096	0.000	0.000	0.000	0.000	0.000	
IV ₂	-1.282	-0.227	0.000	0.000	0.000	0.000	0.000	0.000		IV ₂	-0.964	0.074	0.000	0.000	0.000	0.000	0.000	0.000

Table A7 Non-Zero Modificaiton Indices

Freed parameter			M2	M3
Y_1	ON	Y_2	4.374	2.609
Y_1	ON	IV_2	4.374	2.609
Y_1	ON	Y_1	4.374 ^a	2.609 ^a
Y_2	ON	Y_1	0.745	0.034
Y_2	ON	IV_1	0.745	0.034
Y_2	ON	Y_2	0.745 ^a	0.034 ^a

Note. “ON” refers to regression coefficients

^a Variable regression on itself is a statistical artifact of having structural equations (see text).

Table A8 Covariance Structure for Samples Split by Timea

Group 1 (<1998) ^b										Group 2 (1998+) ^b							
	Y _I	Y ₂	X ₁	X ₂	X ₃	X ₄	IV _I	IV ₂		Y _I	Y ₂	X ₁	X ₂	X ₃	X ₄	IV _I	IV ₂
Means	0.011	21.035	14.189	0.198	8.658	23.262	49.377	0.373		0.059	21.101	15.061	0.125	6.421	27.923	54.221	0.343
Variance	1.058	26.764	3.921	0.045	21.605	14.544	58.611	0.022		0.957	21.993	5.300	0.033	10.128	30.453	36.418	0.025
Standard Deviations	1.029	5.173	1.980	0.212	4.648	3.814	7.656	0.147		0.978	4.690	2.302	0.181	3.182	5.518	6.035	0.158
variable	label	Y _I	Y ₂	X ₁	X ₂	X ₃	X ₄	IV _I	IV ₂	Y _I	Y ₂	X ₁	X ₂	X ₃	X ₄	IV _I	IV ₂
Public Opinion	Y _I	1.000								1.000							
Social Spending	Y ₂	0.563	1.000							0.284	1.000						
Aged	X _I	0.548	0.554	1.000						0.333	0.540	1.000					
Right	X ₂	-0.169	-0.127	0.005	1.000					-0.109	-0.178	0.203	1.000				
Unemp.	X ₃	0.387	0.148	0.014	-0.200	1.000				0.388	0.311	0.233	0.073	1.000			
GDP	X ₄	-0.340	-0.083	0.245	-0.079	-0.622	1.000			-0.481	-0.120	-0.193	-0.016	-0.530	1.000		
FLP	IV _I	-0.657	-0.073	0.018	0.086	-0.521	0.604	1.000		-0.558	-0.059	-0.253	-0.281	-0.667	0.529	1.000	
Veto	IV ₂	-0.180	-0.203	0.094	-0.228	-0.269	0.475	0.072	1.000	-0.046	-0.092	0.196	0.175	0.045	0.314	-0.089	1.000

continued

		Group 1 (<1998) ^b								Group 2 (1998+) ^b							
		Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄	IV ₁	IV ₂	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄	IV ₁	IV ₂
Means		0.011	21.035	14.189	0.198	8.658	23.262	49.377	0.373	0.059	21.101	15.061	0.125	6.421	27.923	54.221	0.343
Variance		1.058	26.764	3.921	0.045	21.605	14.544	58.611	0.022	0.957	21.993	5.300	0.033	10.128	30.453	36.418	0.025
Standard Deviations		1.029	5.173	1.980	0.212	4.648	3.814	7.656	0.147	0.978	4.690	2.302	0.181	3.182	5.518	6.035	0.158
variable	label	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄	IV ₂	IV ₁	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄	IV ₂	IV ₁
Public Opinion	Y ₁	1.058								0.957							
Social Spending	Y ₂	2.998	26.764							1.304	21.993						
Aged	X ₁	1.115	5.676	3.921						0.751	5.833	5.300					
Right	X ₂	-0.037	-0.140	0.002	0.045					-0.019	-0.151	0.085	0.033				
Unemp.	X ₃	1.849	3.566	0.131	-0.197	21.605				1.206	4.638	1.703	0.042	10.128			
GDP	X ₄	-1.332	-1.640	1.849	-0.064	-11.024	14.544			-2.598	-3.109	-2.458	-0.016	-9.306	30.453		
FLP	IV ₁	-5.173	-2.899	0.278	0.140	-18.552	17.644	58.611		-3.295	-1.667	-3.519	-0.307	-12.804	17.622	36.418	
Veto	IV ₂	-0.027	-0.154	0.027	-0.007	-0.183	0.266	0.081	0.022	-0.007	-0.068	0.071	0.005	0.022	0.273	-0.085	0.025

^aSplit by ISSP wave. Group 1: Role of Government (1986, 1990 & 1996), Religion (1991); Group 2: Role of Government (2006), Religion (1998).

^bSee Table 2 for variable coding and country time-points.

Table A9 Testing Equilibrium Comparing Results by Group

var	M2						M3					
	Group 1 (< 1998)			Group 2 (1998 +)			Group 1 (< 1998)			Group 2 (1998 +)		
	b	s.e	β	b	s.e	β	b	s.e	β	b	s.e	β
<i>Y₁</i> ON												
<i>Y₂</i>	-0.010	--	-0.049	-0.010	--	-0.050	0.030	--	0.150	0.030	--	0.146
<i>X₁</i>	0.330	0.047	0.632	0.116	0.055	0.278	0.275	0.040	0.534	0.069	0.054	0.162
<i>X₂</i>	-0.767	0.435	-0.157	-1.675	0.722	-0.315	-0.681	0.374	-0.142	-1.416	0.707	-0.262
<i>X₃</i>	-0.016	0.026	-0.073	-0.027	0.054	-0.089	-0.021	0.022	-0.098	-0.051	0.053	-0.166
<i>X₄</i>	-0.052	0.035	-0.193	-0.039	0.027	-0.226	-0.057	0.031	-0.215	-0.041	0.027	-0.234
<i>IV₁</i>	-0.079	0.014	-0.585	-0.081	0.031	-0.504	-0.076	0.013	-0.569	-0.089	0.030	-0.549
<i>Y₂</i> ON												
<i>Y₁</i>	1.500	--	0.306	1.500	--	0.302	0.750	--	0.150	0.750	--	0.155
<i>X₁</i>	0.895	0.335	0.350	1.122	0.277	0.539	1.154	0.341	0.448	1.201	0.266	0.583
<i>X₂</i>	-1.960	3.213	-0.082	-5.746	3.369	-0.217	-3.206	3.276	-0.133	-6.519	3.244	-0.249
<i>X₃</i>	0.138	0.187	0.126	0.425	0.228	0.282	0.137	0.190	0.125	0.455	0.219	0.305
<i>X₄</i>	0.368	0.246	0.277	0.318	0.140	0.366	0.325	0.251	0.243	0.265	0.134	0.309
<i>IV₁</i>	-7.442	5.128	-0.216	-8.314	4.211	-0.274	-10.944	5.291	-0.315	-7.897	4.043	-0.263
<i>e.Y₁</i>	0.253	0.060	0.243	0.494	0.118	0.547	0.187	0.045	0.185	0.476	0.114	0.512
<i>e.Y₂</i>	13.250	3.231	0.532	11.875	2.839	0.532	13.752	3.290	0.544	11.008	2.632	0.504
RMSEA	0.261			0.170			0.177			0.080		
CFI	0.934			0.943			0.969			0.987		
Exact p	0.034			0.134			0.123			0.295		

Table A10 Implied Covariance Matrices for M3 by Group

Group 1 (< 1998)								
var	Y_1	Y_2	X_1	X_2	X_3	X_4	IV_1	IV_2
Y_1	1.009							
Y_2	2.610	25.275						
X_1	1.085	5.512	3.808					
X_2	-0.036	-0.135	0.002	0.044				
X_3	1.798	3.457	0.125	-0.191	20.987			
X_4	-1.296	1.591	1.797	-0.062	-10.711	14.131		
IV_1	-4.981	-1.630	0.265	0.136	-18.010	17.133	56.940	
IV_2	-0.009	-0.124	0.027	-0.007	-0.179	0.259	0.079	0.021
Group 2 (1998 +)								
var	Y_1	Y_2	X_1	X_2	X_3	X_4	IV_1	IV_2
Y_1	0.928							
Y_2	1.590	21.831						
X_1	0.728	5.768	5.148					
X_2	-0.019	-0.147	0.082	0.032				
X_3	1.173	4.509	1.658	0.041	9.836			
X_4	-2.522	-3.017	-2.382	-0.016	-9.040	29.578		
IV_1	-3.188	-5.018	-3.414	-0.298	-12.443	17.113	35.381	
IV_2	-0.009	-0.066	0.069	0.005	0.022	0.266	-0.082	0.024

Appendix Two and Three

Appendix Two and Three, A2 and A3 available at <https://osf.io/gyz6p>