AMERICAN SOCIOLOGICAL REVIEW

OFFICIAL JOURNAL OF THE AMERICAN SOCIOLOGICAL ASSOCIATION

ONLINE SUPPLEMENT
to article in

AMERICAN SOCIOLOGICAL REVIEW, 2010, VOL. 75 (JUNE:355–377)

The Global Rise of Democracy: A Network Account

Magnus Thor Torfason and Paul Ingram
Columbia Business School – Management Division

SUPPLEMENT 1: THE DYNAMIC NETWORK AUTOREGRESSIVE MODEL

To examine the impact of IGO networks on democracy, we estimate a diffusion model on longitudinal network data. Here, we describe our statistical model and discuss how it relates to other models used for similar purposes. One approach to accounting for diffusion mechanisms is the static network autoregressive model described by Friedkin (1990). Friedkin discusses several variations of this model, the core features of which are:

$$\mathrm{y} = \mathbf{X}\boldsymbol{\beta} + \delta\mathbf{W}\mathrm{y} + \mathrm{u} \quad \text{(S1)}$$

Where y is an n-by-1 vector of dependent variable observations, **X** is an n-by-m matrix containing the observations of the m independent variables, **W** is an n-by-n matrix of network autoregressive effects, u is the error term, and β (m-by-1) and α (scalar) are coefficients to be estimated. The structure of W is not predetermined and takes different forms based on how the network autoregressive effects operate. Because the number of elements (n^2) is larger than the number of observations (n) one must impose an a priori structure on W to estimate the model.

For an unweighted network, one can assume the influence is divided equally between all members in the network. For a weighted network, a natural assumption is that influence is proportional to the strength of the connection. Thus, the influence matrix W can be constructed as simply the network matrix, divided by the number of alters that each member observes (n – 1). For example, in a three-node network (n = 3), the matrix becomes:

$$\mathbf{W} = \frac{1}{2}\begin{bmatrix} 0 & \tau_{12} & \tau_{13} \\ \tau_{21} & 0 & \tau_{23} \\ \tau_{31} & \tau_{32} & 0 \end{bmatrix} \quad \text{(S2)}$$

Here, τ_{ij} represents the strength of the tie between nodes i and j. By transferring all the terms of Equation S1 to the right-hand side of the equality sign, the equation can be rewritten as:

$$\begin{aligned} 0 &= \mathbf{X}\boldsymbol{\beta} + \delta\mathbf{W}\mathrm{y} - \mathrm{y} + \mathrm{u} \\ &= (\delta - 1)\mathrm{y} + \mathbf{X}\boldsymbol{\beta} + \delta\mathbf{W}\mathrm{y} - \delta\mathrm{y} + \mathrm{u} \\ &= (\delta - 1)\mathrm{y} + \mathbf{X}\boldsymbol{\beta} + \delta(\mathbf{W} - \mathbf{I})\mathrm{y} + \mathrm{u} \\ &= \gamma\mathrm{y} + \mathbf{X}\boldsymbol{\beta} + \delta(\mathbf{W} - \mathbf{I})\mathrm{y} + \mathrm{u} \end{aligned} \quad \text{(S3)}$$

If the rows of W are normalized so that the average tie strength in each row is 1 (this is always true in unweighted networks), the value (W – I)y can be rewritten as:

$$\begin{aligned} (\mathbf{W} - \mathbf{I})\mathrm{y} &= \mathbf{K}\mathrm{y} \\ &= \frac{1}{n-1}\begin{bmatrix} -(n-1) & \tau_{12} & \dots & \tau_{1n} \\ \tau_{21} & -(n-1) & & \\ \vdots & & \ddots & \\ \tau_{n1} & & & -(n-1) \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\ &= \frac{1}{n-1}\begin{bmatrix} 0 + \tau_{12}(y_2 - y_1) + \ldots + \tau_{1n}(y_n - y_1) \\ \tau_{21}(y_1 - y_2) + 0 + \ldots + \tau_{2n}(y_n - y_2) \\ \vdots \\ \tau_{n1}(y_1 - y_n) + \tau_{n2}(y_2 - y_n) + \ldots + 0 \end{bmatrix} \end{aligned} \quad \text{(S4)}$$

The individual terms in the last part of Equation S4 represent the network autoregression effect as a multiple of the tie strength τ_{ij} between nodes i and j, and the difference between the values of the dependent variables for the two nodes ($y_j - y_i$). This has an intuitive interpretation as dyadic influence, where each node experiences the influence of each other, proportional to the strength of their tie and the dissimilarity in y between the nodes. Equation S3 shows that in equilibrium, the net effect of the influence of other nodes, the exogenous variables and the individual autoregressive coefficient is zero. In disequilibrium, the effects do not average out, which induces y to move toward the equilibrium state at a rate described by the slope dy, and Equation S3 becomes a model for dynamic network autocorrelation effects:

$$\mathrm{dy} = \gamma\mathrm{y} + \mathbf{X}\boldsymbol{\beta} + \delta\mathbf{K}\mathrm{y} + \mathrm{u} \quad \text{(S5)}$$

Here, **K**y is the vector from Equation S4, representing the net influence of all nodes with which the focal node shares a tie. In discrete time, this translates to:

$$\Delta\mathrm{y}_t = \mathrm{y}_t - \mathrm{y}_{t-1} = \gamma\mathrm{y}_{t-1} + \mathbf{X}_{t-1}\boldsymbol{\beta} + \delta\mathbf{K}_{t-1}\mathrm{y}_{t-1} + \mathrm{u}_t \quad \text{(S6)}$$

There are two main changes between Equations S5 and S6. First, instead of the continuous rate of change (dy), we have the discrete change over one time interval (Δy). Second, each variable is now indexed by time, to reflect the fact that observations from at least two periods are required. This represents the dynamic network autoregressive model (DNAM) for diffusion through one network. The model can be estimated for any panel dataset with network information, and it is the basis of the estimation model we use in the article. In the text, we refer to the $\mathbf{K}_{t-1}y_{t-1}$ term for the IGO network as an "influence variable," denoted as I^{igo}, and incorporate additional terms to control for centrality, other networks, and year-fixed-effects:

$$\Delta y_t = \gamma y_{t-1} + \mathbf{X}_{t-1}\boldsymbol{\beta} + \delta I^{igo}_{t-1} + \rho C^{igo}_{t-1} + \theta I^{ctl}_{t-1} + \phi C^{ctl}_{t-1} + F_t + u_t \qquad \text{(S7)}$$

(In line with other equations in the supplement, S7 is in vector form for a single-period cross-section. In the text, we present this equation in a scalar form for a single observation as Equation 1).

If every network member's influence is the same, as would be the case in a fully connected unweighted network, the model simplifies to a regression to the population mean, consistent with a global isomorphic process. In this case, the matrix **K** is uniformly one, except on the diagonal. Then the vector $\delta \mathbf{K}_{t-1} y_{t-1}$ can be simplified, and for $\kappa = \delta(n-1)/n$ we have:

$$\begin{aligned}\Delta y_t &= \gamma y_{t-1} + \mathbf{X}_{t-1}\beta + \delta \mathbf{K}_{t-1} y_{t-1} + u \\ &= \gamma y_{t-1} + \mathbf{X}_{t-1}\beta + \kappa(\bar{y}_{t-1} - y_{t-1}) + u \end{aligned} \qquad \text{(S8)}$$

Assuming that the focal actor is influenced in the same way by everyone else (completely diffuse effects), a standard model of regression to the population mean gives the same result as DNAM (up to the constant [*n*-1]/n, which is close to 1 for large *n*).

Strang and Tuma (1993) developed the heterogeneous diffusion model (HDM), which is based on a partial likelihood discrete event-history framework. It has some advantages over the static network autoregressive model. Strang and Tuma note that simply including all observations in a static autoregressive model permits later events to influence earlier ones, which is unreasonable if actors have to observe others' behavior to be influenced. This drawback does not affect DNAM, because the only observations affecting the value of the dependent variable at time *t* are the values of the independent variables at time $t-1$.

Strang and Tuma also point out that a regression framework "does not deal naturally with the right censoring generally present in longitudinal data on change in discrete outcomes" (1993:617). This is a problem when examining irrevocable or "final" events, because after adopting a new behavior, an actor will not adopt it again and must be eliminated from the risk set. However, if the relevant behavior is open to reversal or further change, actors should not be removed from the risk set. Such scenarios, which include the one we are examining, are therefore well suited for DNAM, because actors are kept in the risk set even after experiencing a change event.

There are many similarities in the mathematical modeling of HDM and DNAM. A slope of 1 in DNAM indicates that the expected change in the state variable over the next unit of time is 1. Similarly, a hazard rate of 1 in HDM indicates that the expected change in the state variable is 1. Thus, Equation S7 (which describes DNAM) is in many ways comparable with the following equation, which describes HDM:

$$r_n(t) = \exp(\alpha' \mathrm{x}_n) + \exp(\beta' \mathrm{v}_n) \sum_{s \in \mathscr{S}} \exp[\gamma' \mathrm{w}_s + \delta' z_{ns} + \zeta(t - t_s)]$$
$$= \exp(\alpha' \mathrm{x}_n) + \exp(\beta' \mathrm{v}_n) \sum_{s \in \mathscr{S} \cup \mathscr{N}} (y_s - y_n) \exp(\gamma' \mathrm{w}_s) \exp(\delta' z_{ns}) \exp(\zeta(t - t_s))$$

(S9)

Here, variables in x affect intrinsic propensity to change (comparable to the matrix X in Equation S7), variables in v affect the susceptibility of the focal actor to influence, variables in w affect the infectiousness of the alter, and variables in z reflect the network structure. The way HDM models influence as the sum over only the actors in S (prior adopters) is algebraically identical to summing over all alters while interacting each term with the difference between the alter and the focal actor ($y_s - y_n$). Because y is 1 for adopters and 0 for non-adopters, this difference is 0 for two non-adopters, and their inclusion in the sum is irrelevant.

The estimated parameters are not fully comparable, due to the exponents in HDM. Nevertheless, the use of interactions between different types of parameters closely mirrors the use of such interactions in DNAM. Because Strang and Tuma use Maximum Likelihood to estimate their model, they include all effects in a single interaction and omit all partial interactions. In DNAM, which is estimated using a linear or probit regression model, each interaction is included separately, and all partial interactions must be included to avoid omitted variable bias. While there are certain differences between these approaches, the philosophy is similar. The tendency to change is presumed to be governed by intrinsic properties, but also by diffusion effects, which are assumed to impact the propensity to change, and the choice of method depends mainly on the form of data in question.

SUPPLEMENT 2: AUXILIARY STATISTICAL ANALYSIS

Due to space considerations, not all of our analysis could be presented in the main text of the article; this supplement covers some of the omitted material.

Table S1 shows descriptive statistics and correlations for the variables we use in our analysis. Some variables have been standardized to aid in interpreting coefficients.

Table S2 presents a number of auxiliary estimations, which we performed to establish the robustness of the results we present in the article. Model MS-1 is simply a reproduction of Model 4 in the text, repeated here for reference. The following models are presented as variations of this model to aid in comparison.

In MS-2 through MS-5, we incorporate different dummy variables and fixed-effect specifications. We examine specifications with dummy variables for continents (MS-2), periods (instead of using year-fixed-effects [MS-3]), and both continents and periods (MS-4). The estimate for diffusion through IGOs is very stable across specifications, and the models concur regarding other effects as well. Compared with Europe (the omitted continent), the continental effect is significant and negative for Africa, Asia, and the Americas (North, Central, and South America are grouped together). Compared to 1989 to 2000 (the omitted period), the period effect is significant and negative for 1919 to 1938 (the Age of Dictators) and 1946 to 1988 (the Cold War). MS-5 includes state-fixed-effects in addition to year-fixed-effects, and our results are robust to this alternative specification. In analysis not included in the table, we estimated our models with certain regions or continents omitted. For example, we omitted the countries involved with the Soviet Union and the Warsaw Pact, as well as the states involved with the European Union, to ensure the effect does not stem from particular idiosyncrasies of these blocs, but this does not affect our results. We also restricted our analysis to particular time periods. In almost all cases, both the size and significance level of the main effect are maintained. When including observations only from the eighteenth century—the infancy of the IGO system—the size of the coefficient is maintained, but it drops out of significance because the standard errors increase. This is most likely due to the large number of states whose ties to the international system were extremely marginal. When we run our analysis of the eighteenth century only for states that were members in five or more international organizations, the effect is highly significant ($p < .01$).

In MS-6 through MS-11, we estimate several specifications where we restrict or omit variables or observations to examine the robustness of our results. MS-6 omits the lagged dependent variable; MS-7 omits the influence variable for global diffusion; MS-8 omits the influence variable for global diffusion and the measure of centrality in the IGO network; and MS-9 omits *all* network level control variables. MS-10 includes a “minimalist” specification, omitting all fixed effects and including only lagged democracy and the two factors of the interaction we use to construct our main variable (diffusion through the fully connected network and the degree centrality score in the IGO network). MS-11 omits all observations for “new” states that have been in the system for fewer than 10 years. The coefficient for IGO diffusion is robust to all these specifications, and the point estimate lies between the basic estimate presented in Model 4 and the main robustness analyses presented in Models 15 and 16.

MS-12 through MS-15 incorporates control variables that are theoretically relevant, but for which even the best sources known to us contain a considerable amount of missing observations for the period and scope we are examining. In choosing which variables to incorporate, a balance must be struck between the potential for omitted variable bias and selection bias. By presenting these estimates here, we attempt to provide reasonable reassurance against both.[1]

MS-12 examines the effect of education, specifically university enrollment. We find a positive and significant effect for this variable. MS-13 examines the effect of exposure to world society through memberships in INGOs. We find a positive and significant effect, in line with the predictions of world polity theorists. MS-14 estimates the impact of

[1] In the main text of the article, we discuss the relevant data sources and provide references.

outward FDI (when the focal country has made investments abroad) and inward FDI (when the focal country has received investment). We do not find significant effects for either measure. Finally, MS-15 examines the network of bilateral trade. In the main text of the article, we incorporate the trade network only when examining the post-1950 period, but here we examine its impact using bilateral trade data from Barbieri (1998), which goes back to 1850 (these data are missing a considerable number of observations). The results are similar to what we find for recent years: there is no significant effect of the trade network.

In all specifications, the key effect for our theory—diffusion of democracy through the IGO network—is robust in both size and significance levels. Because the missing observations for the four variables examined here are not completely overlapping, incorporating them all in one model forces us to throw away around 90 percent of our observations. We therefore do not include the model in the table; note, though, that the coefficient for IGO diffusion maintains its size and significance in that specification.

Table S3 reports results of a propensity score matching procedure, where states that experience a significant positive democratization influence are compared directly with states that do not, but whose propensity to experience such influence based on their other characteristics is similar. When we look at states that experience a large positive influence, we see that on average they experience a .50 annual increase in their democracy level, compared with no increase on average for the control group that did not experience such influence. The matched control group experiences an average annual increase of .23, but even when comparing matched observations, the difference (.27 per year) is significant ($p < .015$).

Figures S1 and S2 show estimation coefficients that could not be displayed (and would be hard to interpret) in tabular form. The figures show coefficients as estimated from Model 4. Figure S1 shows the values of the year fixed effects. Consistent with the period effects estimations in models MS-3 and MS-4, the figure shows a drop from 1919 through 1988.

Our estimations use ordered-probit, which does not rely on the dependent variable being an interval measure (as OLS does). Rather, it estimates a set of “cut-points” that reflect the relationship between the effect of the independent variables and the actual outcomes in the dependent variables. This counters issues that would otherwise arise, both because of floor/ceiling effects (due to the bounded range of the democracy measure), and because structural change may be “clumpy” rather than gradual over time. Figure S2 shows the cut-point values resulting from the ordered-probit regression. The large gap in the center demonstrates that the most likely outcome for any state in any given year is no change in democracy. However, when the combination of covariates and the “error draw” is so large that change occurs, the change can be large or small. This corresponds to what we see in the data and to what we know about institutional change: structures often remain stable for a long time, but when change occurs, multiple structures may change at the same time, sometimes in idiosyncratic ways.

TABLE S1. Summary Statistics

		Mean	S.D.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(18)	(19)	(20)
(1)	Democracy	-0.7	7.1																			
(2)	Per-capita Iron/Steel Production	0.1	0.8	.31																		
(3)	Per-capita Energy Consumption	1.7	5.3	.02	.10																	
(4)	CINC Score	0.0	0.0	.09	.02	.04																
(5)	Logged Per-capita GDP	7.7	1.3	.44	.17	.39	.06															
(6)	Total IGO Membership	28.2	24.2	.36	.14	.20	-.10	.56														
(7)	Centrality in IGO Network	1.0	1.0	.24	.09	.19	-.16	.58	.92													
(8)	Centrality in Distance Network	1.6	1.0	.04	.13	.17	-.14	.23	.45	.51												
(9)	Centrality in Continental Network	2.0	1.0	-.03	.02	.14	-.17	-.03	.58	.71	.60											
(10)	Centrality in Regional Network	1.6	1.0	-.04	-.04	.12	-.30	-.06	.41	.53	.42	.70										
(11)	Centrality in Alliance Network	0.7	1.0	.01	.07	.30	-.01	.19	.52	.46	.23	.24	.26									
(12)	Centrality in Colonial Network	0.3	1.0	.16	.01	.03	-.00	.14	.28	.19	.08	.07	.04	.11								
(13)	Centrality in Trade Network	0.8	1.0	.33	.42	.24	.14	.62	.78	.68	.31	.23	.00	.31	.35							
(14)	Global Diffusion	0.0	1.0	-.87	-.32	.02	-.09	-.36	-.26	-.12	.03	.20	.17	.01	-.17	-.41						
(15)	IGO Network Diffusion	0.0	1.0	-.74	-.33	-.00	-.07	-.38	-.22	-.11	.03	.21	.18	.02	-.19	-.43	.94					
(16)	Distance Network Diffusion	0.0	1.0	-.81	-.23	.03	-.08	-.28	-.21	-.11	-.01	.10	.08	.09	-.15	-.33	.88	.82				
(17)	Continental Network Diffusion	0.0	1.0	-.76	-.14	.08	-.11	-.11	-.13	-.07	.05	-.00	.02	.11	-.08	-.21	.81	.72	.83			
(18)	Regional Network Diffusion	0.0	1.0	-.60	-.02	.14	-.01	-.06	-.07	-.05	-.01	.00	.00	.10	-.03	-.07	.65	.58	.69			
(19)	Alliance Network Diffusion	0.0	1.0	-.39	-.11	.02	-.16	-.03	-.05	-.02	-.02	.00	.03	-.07	-.07	-.12	.42	.40	.38	.45		
(20)	Colonial Network Diffusion	0.0	1.0	-.24	-.23	-.07	-.06	-.22	-.21	-.08	-.05	.10	.09	-.14	-.78	-.39	.29	.31	.25	.08	.09	
(21)	Trade Network Diffusion	0.0	1.0	-.64	-.46	-.02	-.19	-.38	-.34	-.15	-.02	.18	.24	-.00	-.34	-.57	.75	.82	.67	.33	.32	.45

ONLINE SUPPLEMENT
to article in
AMERICAN SOCIOLOGICAL REVIEW, 2010, VOL. 75 (JUNE:355–377)

TABLE S2. Auxiliary Regression Results

	MS-1: Model 4	MS-2: Continental effects	MS-3: Period effects	MS-4: Continental & Period effects	MS-5: State/Year Fixed Effects	MS-6: LDV omitted	MS-7: Global diffusion omitted	MS-8: Global diff. & Centr. omitted	MS-9: All network controls omitted	MS-10: Minimalist specification	MS-11: New states omitted	MS-12: University enrollment	MS-13: INGO memberships	MS-14: FDI	MS-15: Bilateral trade
Democracy (Lagged)	-.016*	-.019*	-.022**	-.025***	-.034*		-.003	-.001	-.012**	-0.019***	-.026**	-.019	.001	.115	-.014
	(.008)	(.008)	(.007)	(.007)	(.013)		(.005)	(.005)	(.004)	(0.006)	(.009)	(.010)	(.054)	(.082)	(.012)
Per capita Iron/Steel Production	.161	.112	.191	.163	.119	.143	.143	.164	.347*		.215	.105	.069	.271	.225
	(.150)	(.142)	(.148)	(.145)	(.171)	(.147)	(.150)	(.149)	(.136)		(.157)	(.153)	(.135)	(.187)	(.223)
Per capita Energy Consumption	-.004	-.004	-.004	-.004	.010	-.005	-.005	-.005	-.004		-.010	-.003	-.003	-.015*	-.001
	(.003)	(.003)	(.003)	(.003)	(.008)	(.003)	(.003)	(.003)	(.003)		(.005)	(.003)	(.003)	(.006)	(.003)
CINC Score	.676	.572	.501	.329	-.010	.602	.594	.637	.331		.964	.630	.761	-.539	.849
	(.459)	(.414)	(.452)	(.415)	(.639)	(.440)	(.453)	(.457)	(.402)		(.498)	(.565)	(1.168)	(1.635)	(.580)
Total IGO Membership	.011**	.009*	.011***	.009**	-.000	.011**	.012**	.008***	.007***		.016***	.010*	.004	-.003	.014*
	(.004)	(.004)	(.003)	(.003)	(.007)	(.004)	(.004)	(.002)	(.002)		(.005)	(.004)	(.005)	(.006)	(.006)
Global Diffusion	-.220	-.131	-.266**	-.183	-.195	-.071				-0.169*	-.352*	-.181	-.189	.468	-.228
	(.119)	(.128)	(.103)	(.117)	(.185)	(.076)				(0.071)	(.145)	(.137)	(.344)	(.507)	(.173)
Centrality in IGO Network	-.193	-.095	-.099	-.043	-.008	-.148	-.178			0.110***	-.507**	-.139	-.184	.117	-.282
	(.156)	(.154)	(.081)	(.087)	(.301)	(.156)	(.157)			(0.018)	(.176)	(.172)	(.194)	(.223)	(.228)
IGO Network Diffusion	.243***	.240***	.249***	.239***	.303**	.172***	.135***	.133***	.158***	0.230***	.327***	.240**	.272**	.245*	.352***
	(.068)	(.069)	(.060)	(.063)	(.094)	(.051)	(.037)	(.037)	(.027)	(0.048)	(.081)	(.077)	(.084)	(.112)	(.106)
Centrality in Distance Network	-.012	-.027	-.024	-.036	.012	-.013	-.010	-.011			.007	-.017	-.023	.010	-.023
	(.024)	(.020)	(.021)	(.019)	(.108)	(.024)	(.024)	(.024)			(.024)	(.027)	(.021)	(.024)	(.027)
Distance Network Diffusion	.016	.025	.013	.017	.034	.028	.009	.007			-.003	.017	-.034	-.040	-.004
	(.044)	(.044)	(.044)	(.045)	(.061)	(.044)	(.043)	(.042)			(.046)	(.053)	(.048)	(.060)	(.054)
Centrality in Continental Network	-.057	-.112*	-.024	-.045	-.106	-.058	-.078	-.070			-.070	-.101*	.042	-.004	-.120
	(.044)	(.050)	(.045)	(.049)	(.073)	(.043)	(.042)	(.044)			(.049)	(.050)	(.064)	(.075)	(.062)
Continental Network Diffusion	-.022	-.123	-.025	-.103	.023	-.013	-.043	-.033			-.030	-.060	.097	-.027	-.016
	(.051)	(.065)	(.051)	(.065)	(.085)	(.051)	(.053)	(.052)			(.053)	(.060)	(.071)	(.074)	(.071)
Centrality in Regional Network	.011	.019	.015	.025	-.046	.008	.008	.005			.033	.027	-.012	-.016	.020
	(.028)	(.030)	(.028)	(.031)	(.070)	(.028)	(.027)	(.028)			(.029)	(.031)	(.034)	(.041)	(.037)
Regional Network Diffusion	.119**	.127***	.117**	.124***	.124*	.115**	.118**	.119**			.113**	.118**	.052	.142**	.074
	(.037)	(.037)	(.037)	(.037)	(.052)	(.037)	(.037)	(.037)			(.036)	(.041)	(.048)	(.051)	(.044)
Centrality in Alliance Network	-.028	-.021	-.016	-.006	-.025	-.027	-.025	-.022			-.037	-.040	-.021	.008	-.028
	(.026)	(.029)	(.026)	(.028)	(.038)	(.026)	(.026)	(.027)			(.026)	(.029)	(.026)	(.033)	(.030)
Alliance Network Diffusion	.067**	.056*	.063**	.055*	.055*	.069**	.066**	.066**			.077**	.060*	.082**	.071*	.054*
	(.023)	(.024)	(.023)	(.023)	(.027)	(.024)	(.023)	(.023)			(.024)	(.024)	(.026)	(.035)	(.025)
Centrality in Colonial Network	.038	.043*	.036	.040*	-.016	.039	.038	.039			.044*	.042	.045	-.001	.055
	(.023)	(.019)	(.020)	(.019)	(.026)	(.023)	(.023)	(.024)			(.022)	(.023)	(.033)	(.030)	(.028)
Colonial Network Diffusion	.049*	.057*	.050*	.057*	-.007	.051*	.052*	.048*			.057**	.045*	.039	-.004	.054*
	(.023)	(.022)	(.022)	(.023)	(.025)	(.024)	(.023)	(.024)			(.022)	(.022)	(.030)	(.036)	(.024)
Continent = Africa		-.267**		-.215**											
		(.087)		(.082)											
Continent = Americas		-.146*		-.118*											
		(.064)		(.059)											
Continent = Asia		-.208*		-.137											
		(.093)		(.087)											
Continent = Oceania		-.198		.000											
		(.117)		(.099)											
1815 to 1913			.012	-.030											
			(.130)	(.141)											
1814 to 1918			.075	.051											
			(.157)	(.163)											
1919 to 1938			-.406**	-.424**											
			(.126)	(.133)											
1939 to 1945			-.239	-.257											
			(.166)	(.166)											
1946 to 1988			-.312***	-.296***											
			(.086)	(.085)											
Logged University Enrollment per capita												.047*			
												(.020)			
Total INGO Membership													.118**		
													(.040)		
Stock of Inward FDI														-.010	
														(.017)	
Stock of Outward FDI														.018	
														(.017)	
Centrality in Trade Network															-.036
															(.046)
Trade Network Diffusion															-.042
															(.042)
Observations	10633	10633	10633	10633	10633	10633	10633	10633	11086	11086	9234	8904	4524	2314	6997
R-squared															

* $p < .05$; ** $p < .01$; *** $p < .001$

TABLE S3. Propensity Score Matching

Outcome on treated group:	.50
Outcome on untreated (unmatched):	–.03
Outcome on untreated (matched):	.23
Difference (matched):	**.27**
Standard error (Bootstrap):	.11
P value:	**.014**

Note: The table shows the results of estimating the effect of democratization influence through the IGO network on democratic change, after matching for covariates likely to influence the propensity for treatment (the covariates in Model 4 are used for matching). Specifically, the treatment is “experiencing positive democratization influence more than one standard deviation above mean,” and outcome is “change in observed democracy level.” *P* value denotes significance of the difference between the outcome on the treated group and the untreated, matched sample.

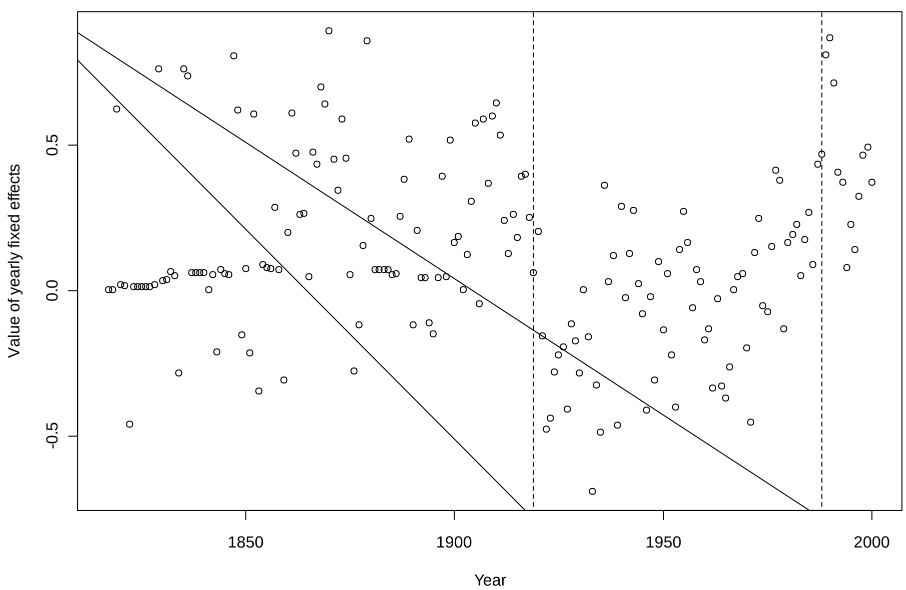

FIGURE S1. Value of Yearly Fixed Effects
Note: The figure shows the value of the year-fixed-effects, as estimated in Model 4. A high value indicates an unusually high tendency for states to experience positive democratic change, and vice versa. The values vary quite a bit per year, but they show a sharp drop at 1919, coinciding with the start of the age of dictators. After the drop, they edge up again and are quite high around and after 1988. The dotted lines demarcate the 1919 to 1988 period.

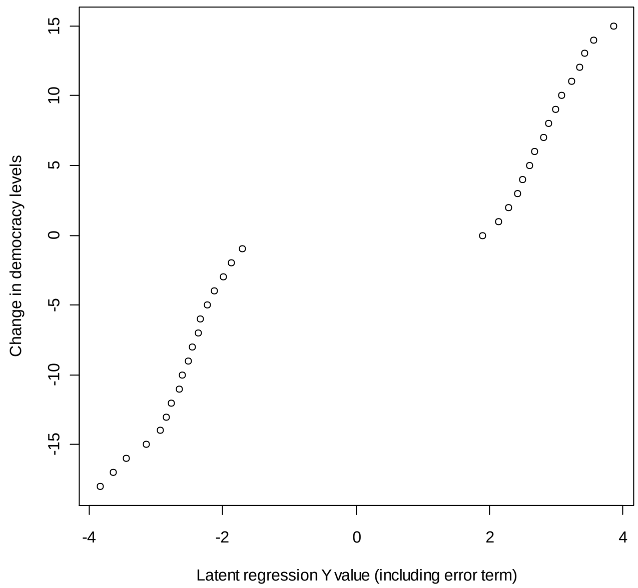

FIGURE S2. Ordered Probit Regression Cut-Point Values
Note: The figure shows ordered probit cut-point values, as estimated in Model 4. These values form a (multiple) step function, which allows a mapping from the latent regression underlying the model to the observed change in democracy value. For each observation, the process is modeled as being determined by a Y value in the latent regression, determined by the independent variables, as well as by a normally distributed error term. The step function is used to map this Y value into the resulting change in democracy levels. If the mapping was linear (as is assumed when using OLS), this would be close to a straight line. However, as can be seen from the graph, the widest range of latent Y values are mapped a zero change in democracy levels. This is consistent with the data; by far, the most common outcome is no change in democracy levels.

References

Barbieri, Katherine. 1998. "Notes on International Trade Database." *The Correlates of War Project,* (http://www.correlatesofwar.org/).

Friedkin, Noah E. 1990. "Social Networks in Structural Equation Models." *Social Psychology Quarterly* 53:316–28.

Strang, David and Nancy Brandon Tuma. 1993. "Spatial and Temporal Heterogeneity in Diffusion." *American Journal of Sociology* 99:614–39.