Supplementary Materials for Including Measurement Uncertainty in Time-Series, Cross-Sectional Analyses: The Case of Support for Democracy

Contents

List of Figures

List of Tables

1. Conceptualization of Democratic Mood

Democratic mood captures the extent to which a national public offers explicit support for a democratic system and rejects any autocratic alternatives (Linz and Stepan 1996). It is principled or diffuse support for democracy itself, rather than instrumental support for the outputs of government or the incumbent office-holders. Consequently, mood is measured using existing survey questions which ask respondents to evaluate the appropriateness or desirability of democracy; compare democracy to some undemocratic alternative; or evaluate one of these undemocratic forms of government. Such items are widely used to measure democratic support (e.g., Bratton, Mattes, and Gyimah-Boadi 2005; Magalhães 2014; Norris 2011). Questions focusing on related concepts such as satisfaction with the performance of democracy and trust in national political institutions were not included because neither is a valid measure of principled support for democracy (e.g., Canache, Mondak, and Seligson 2001; Linde and Ekman 2003). A full list of the survey items which were used are available in the replication materials folder, here (url to be added).

2. Modeling Approach

2.1. The dynamic latent variable model

Claassen (2019) develops six versions of a dynamic latent variable model for estimating TSCS national opinion, recommending the fifth and sixth of these. While both perform well, the fifth is simpler and is used to measure support for democracy in Claassen (2020*a*) and Claassen (2020*b*). It employs a beta-binomial link function between observed, nationally-aggregated survey responses y_{ikt} for each country i, year t, and survey questions k, and the probabilities of offering a supportive response π_{ikt}. These probabilities are modeled as a function of the latent country-year opinion estimates of interest θ_{it}, item bias parameters λ_k and item-country bias parameters δ_{ik}. The latent estimates are furthermore modeled as evolving over time via a random walk process:

$$y_{ikt} \sim \text{Beta-Binomial}(s_{ikt},\ \pi_{ikt},\ \phi) \tag{1}$$

$$\pi_{ikt} = \text{logit}^{-1}(\delta_{ik} + \lambda_k + \theta_{it}) \tag{2}$$

$$\theta_{it} \sim \text{N}(\theta_{i,t-1},\ \sigma^2_{\theta}) \tag{3}$$

Claassen's (2019) more complex sixth model is used in the present paper. It adds item slopes or discrimination parameters γ_k to the above model. Specifically, equation (2) above is expanded as follows:

$$\pi_{ikt} = \text{logit}^{-1}(\delta_{ik} + \lambda_k + \gamma_k\theta_{it}) \tag{4}$$

These item slopes allow analysts to test if survey questions fit the single dimension of opinion that is assumed to underlie the observed survey responses. It is therefore recommended for general use and is incorporated into the unified model developed below. To align the analyses conducted in this paper, mood is re-estimated using the sixth model. These estimates are overwhelmingly similar (Pearson's correlation = 0.99).

2.2. Structural models of mood and democracy

Claassen (2020*a*) models democracy d in any given country i and year t as a function of its previous two lags, lagged mood m, and several covariates B. The parameter δ_1 captures the effect of mood on subsequent democracy, i.e., the Lipset hypothesis.

$$d_{it} = \mu^{(d)} + \zeta_1^{(d)} d_{it-1} + \zeta_2^{(d)} d_{it-2} + \delta_1 m_{it-1} + X_{it-1}^{(m)} \mathrm{B}^{(d)} + \epsilon_{it}^{(d)} \tag{5}$$

Claassen (2020*b*) then models mood m in any given country-year as a function of its previous two lags, the first lag and first difference of democracy, and other covariates. The parameter δ_2 captures the effect of democracy on subsequent mood, i.e., the socialization hypothesis, while δ_3 captures the immediate effect of change in democracy on mood, i.e., the thermostatic hypothesis.

$$m_{it} = \mu^{(m)} + \zeta_1^{(m)} m_{it-1} + \zeta_2^{(m)} m_{it-2} + \delta_2 d_{it-1} + \delta_3 \Delta d_{it} + X_{it-1}^{(m)} \mathrm{B}^{(m)} + \epsilon_{it}^{(m)} \tag{6}$$

2.3. Incorporating measurement uncertainty via the method of composition

Tai, Hu, and Solt (2022) adopt a two-step approach to including measurement uncertainty, described initially by Tanner (1996) as the "method of composition" and introduced to political science by Treier and Jackman (2008). In a situation with a country and year-varying latent variable θ_{it}, the analyst takes a number of draws p from the posterior distribution of their latent estimates $\tilde{\theta}_{pit}$. These draws, rather than than a single set of point estimates, are then employed in subsequent analyses. We illustrate with the following model of mood, which incorporates draws from the posterior distributions of democracy $\tilde{\theta}^{(d)}$ and mood $\tilde{\theta}^{(m)}$:

$$\tilde{\theta}_{pit}^{(m)} = \mu^{(m)} + \zeta_1^{(m)} \tilde{\theta}_{pit-1}^{(m)} + \zeta_2^{(m)} \tilde{\theta}_{pit-2}^{(m)} + \delta_2 \tilde{\theta}_{it-1}^{(d)} + \delta_3 \Delta \tilde{\theta}_{it}^{(d)} + X_{it-1} \mathrm{B}^{(m)} + \epsilon_{it}^{(m)} \tag{7}$$

Such a model is fit p times, once for each posterior draw of mood and democracy. This captures measurement uncertainty. Parameter uncertainty is then included by simulating a draw from the j parameter by p draw matrix of coefficient estimates B and the p-length array of j by j by variance-covariance matrices Σ (which may be robust or conventional variance-covariance matrices). Point estimates, standard errors, and confidence intervals can be extracted from the resulting vector of draws of estimates for each parameter: the mean of the vector can be used as the point estimate, with the standard deviation providing an estimate of the standard error.

2.4. The unified model

An alternative approach to handling measurement error, proposed in the present paper, is to incorporate both the measurement and structural models in a single likelihood function. The analyst jointly estimates latent variables and the structural links between latent and observed variables. Such a joint, or unified model has appeared in a number of guises (e.g., Kellstedt, McAvoy, and Stimson 1993/94; Skrondal and Rabe-Hesketh 2004).

We describe our unified model below. The model begins in the same fashion as the dynamic latent variable model presented above. However the third line now integrates the structural model of mood (equation 6 above) and the dynamic model of latent opinion (equation 3 above):

$$y_{ikt} \sim \text{Beta-Binomial}(s_{ikt},\ \pi_{ikt},\ \phi) \tag{8}$$
$$\pi_{ikt} = \text{logit}^{-1}(\delta_{ik} + \lambda_k + \gamma_k \theta_{it}^{(m)}) \tag{9}$$
$$\theta_{it}^{(m)} = \mu^{(m)} + \zeta_1^{(m)} \theta_{it-1}^{(m)} + \zeta_2^{(m)} \theta_{it-2}^{(m)} + \delta_2 \theta_{it}^{(d)} + \delta_3 \Delta\theta_{it}^{(d)} + X_{it-1}^{(m)} \mathrm{B}^{(m)} + \epsilon_{it}^{(m)} \tag{10}$$

Democracy has a simpler measurement model because annual point estimates and standard deviations are available from V-Dem for every country. Observed democracy scores d^{obs} are treated as a function of an unobserved, "true" democracy score $\theta^{(d)}$, with the degree of error measured using the observed standard deviations in annual democracy scores d^{sd}. These latent estimates, rather than the observed V-Dem scores, are then used in the structural model of democracy:

$$d_{it}^{obs} \sim N(\theta_{it}^{(d)}, d_{it}^{sd}) \tag{11}$$
$$\theta_{it}^{(d)} = \mu^{(d)} + \zeta_1^{(d)} \theta_{it-1}^{(d)} + \zeta_2^{(d)} \theta_{it-2}^{(d)} + \delta_1 \theta_{it-1}^{(m)} + X_{it-1}^{(d)} \mathrm{B}^{(d)} + \epsilon_{it}^{(d)} \tag{12}$$

All of these steps, equations 8 through 12, are estimated simultaneously in the unified model.

3. Model Specification and Estimation

We fit three models using Bayesian MCMC methods: (1) the univariate latent variable model described in section 2.1 above and originally proposed and used by Claassen (2019; 2020*a*;*b*); (2) the unified model that estimates an overall effect of mood on democracy; (3) the unified model which allows this effect of mood on democracy to vary by regime. Each of these models is employed twice – once using Claassen's original data and again using an expanded dataset. The estimation of these models is described here.

3.1. Specifying the univariate latent variable models

We begin by using the dynamic latent variable described in section 2.1 to estimate mood. We include several computational refinements over the model developed by Claassen (2019). First, we allow for ragged country-by-year arrays to accommodate the varying length of national latent opinion time-series (due to the varying years in which survey measurement commenced). We also make use of non-centered parameterizations for all variance terms, e.g., $\sigma_{\theta^{(m)}}$, $\sigma_{\gamma^{(m)}}$, and $\sigma_{\delta^{(m)}}$. Non-centered parameterizations include standard-normally distributed redundant parameters, e.g., $\nu_{ik}^{\delta^{(m)}}$ which shift variance and covariance terms away from zero, making MCMC sampling more efficient:

$$\delta_{ik}^{(m)} = \sigma_{\delta^{(m)}}^2 \times \nu_{ik}^{\delta^{(m)}} \tag{13}$$

The item-country variances are given weakly-informative half-Normal priors, e.g., $\sigma_\delta^{(m)} \sim \mathrm{N}^+(0,\ 1)$. The variance-covariance matrix for the item intercepts λ and slopes γ is split into two variances and correlation term, with the former receiving a half-Normal $(0, 1)$ prior and the latter an LKJ (2) prior. Item intercepts and slopes are identified by setting their expectations: the former at the log of the mean proportion expressing support for democracy, and the latter at 0.5. The beta-binomial dispersion parameter ϕ receives a gamma$(3,\ 0.04)$ prior. Since latent opinion is modeled as a function of its value in the previous year, we estimate initial values for each country, in the year preceding the first estimates based on data. These initial values receive a N(0, 1) prior.

3.2. Specifying the unified models

Non-centered parameterizations are again used for all variance terms. Variance parameters all receive half-normal $N^{+}(0, 1)$ priors. Parameters capturing the effects of lagged outcome variables are restricted to ensure stationarity of each time series. The second lag is restricted to lie between -1 and +1 and is given a Uniform $(-1, 1)$ prior. The sum of first and second lags is restricted to lie between 0 and 1; it receives a weakly informative Beta $(3, 1)$ prior. The first lag is then defined as the difference between the lag sum and the second lag.

Since the models of mood and democracy require two lags of each, we estimate two years' worth of initial values for mood and democracy. These initial values are given N(0, 1) priors. Regression / structural parameters are given N(0, 1) priors. Structural models only take the initial values of democracy or mood as inputs into lagged outcomes (and necessarily only in years 1 and 2). The values used as outcome variables are based only on estimates obtained for years in which mood survey data are available. Structural model residuals are also given a non-centered parameterization.

3.3. Estimation

These Bayesian models are estimated with Bayesian Markov-Chain Monte Carlo (MCMC) methods. Stan software, which implements Hamiltonian Monte Carlo sampling (Carpenter et al. 2017), is employed. Four parallel chains, with randomly selected starting values drawn from a Uniform $(-1, 1)$ distribution, are run, with 500 warmup and 1,500 post-warmup samples each. The 4,000 post-warmup samples are saved and analyzed further.

4. Model Checking

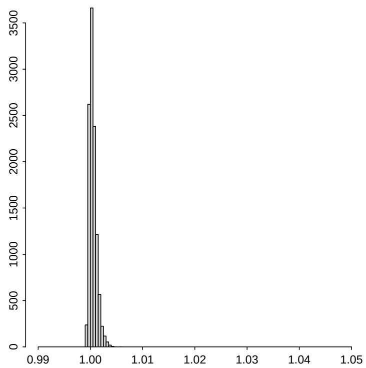
Convergence is assessed using a variety of diagnostics, including traceplots of multiple parameters and Gelman-Rubin R-hat statistics. The latter were close to one for all models (Figure S1), indicative of convergence.

The models can be further verified using posterior predictive checking: simulating data conditional on the estimated parameters and comparing the simulated data against the actual data used to produce the estimated parameters (Gelman et al. 2014). As Figure S2 shows, there is a close correspondence between the aggregated survey responses for each of the national mood items in our dataset $y_{ikt}^{(m)}$ (we show these here as proportions, i.e., $y_{ikt}^{(m)}/s_{ikt}^{(m)}$, for ease of presentation) and those we simulate $\tilde{y}_{ikt}^{(m)}$ in each of our six Bayesian models.[1] These posterior predictive checks suggest that each model fits the observed survey responses.

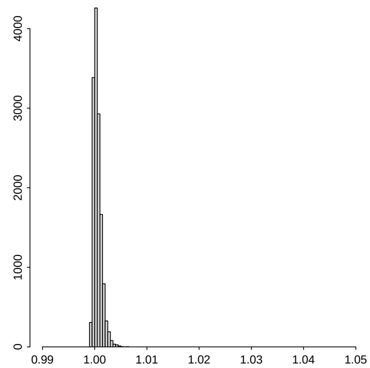
[1] These models are the measurement model to generate estimates for MEX and MOC analyses and two unified models, with either overall effects of mood, or regime-specific effects. Each of these three is run using the original and extended datasets.

Figure S1. Model convergence as shown by the Gelman-Rubin R-hat diagnostic

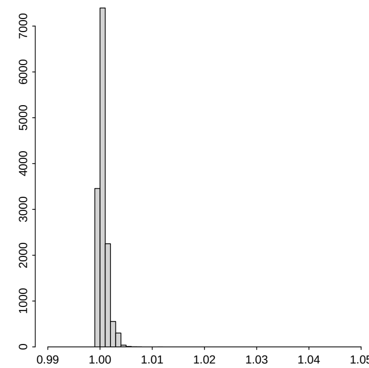
(a) Measurement model, original data



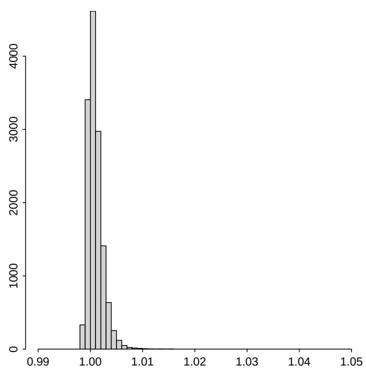
(b) Unified model, overall mood effect, original data



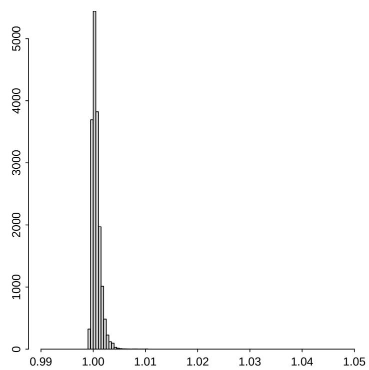
(c) Unified model, regime-specific mood effect, original data



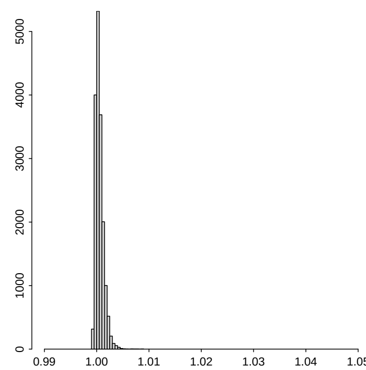
(d) Measurement model, extended data



(e) Unified model, overall mood effect, extended data



(f) Unified model, regime-specific-mood effect, extended data



Distribution of the Gelman-Rubin R-hat statistic for all parameters, one plot for each model.

5. Data

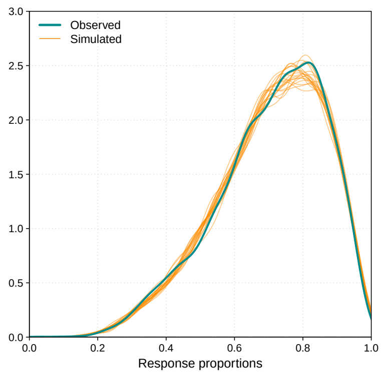
5.1. Mood

Two datasets are used to estimate democratic mood. The first is the original dataset collected by Claassen (2020*a*; 2020*b*), which runs from 1988 (in some cases) to 2017. Following Claassen's initial coding rules, countries with less than two years' of survey measures were dropped, as were countries lacking V-Dem democracy data (either non-independent territories or micro-states). Data from items that were fielded in only one wave of surveys were also dropped to ensure that item-country parameters were identifiable. There are 3,531 nationally aggregated survey responses remaining, drawn from 13 survey projects and 135 countries.

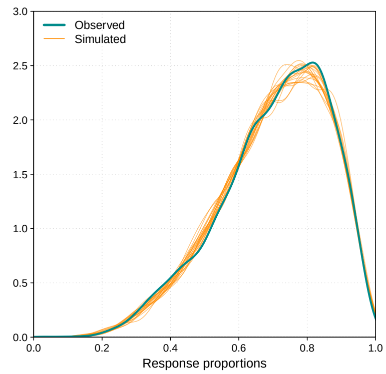
The models were also run on an expanded dataset. Although Tai, Hu, and Solt (2022) collect and publish their own expanded dataset, it includes a number of errors and inconsistencies (see next section), necessitating that this data be recollected. This is accomplished by extending Claassen's original dataset to 2020 by adding new survey measures from the survey projects that

Figure S2. Posterior predictive plots

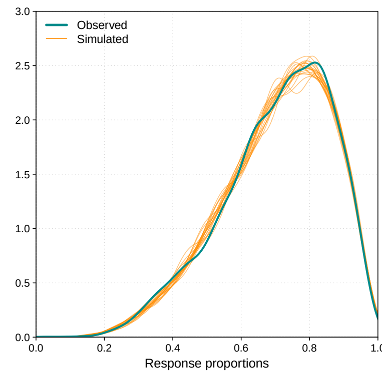
(a) Measurement model, original data



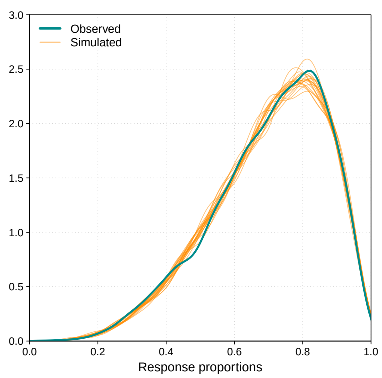
(b) Unified model, overall mood effect, original data



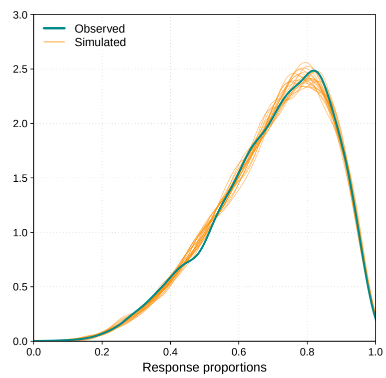
(c) Unified model, regime-specific mood effect, original data



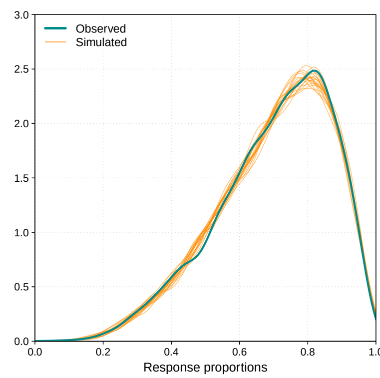
(d) Measurement model, extended data



(e) Unified model, overall mood effect, extended data



(f) Unified model, regime-specific-mood effect, extended data



Each plot compares the distributions of response proportions in the observed survey measures of mood (teal line) against 20 draws from the posterior distributions of the response proportions π_{ikt} estimated by the particular model (orange lines).

were originally included. Data from two smaller cross-national survey projects – the Arab Transformation Project and the Consolidation of Democracy in Central and Eastern Europe – were also added. Once again, countries with less than two years' data are dropped. Survey items fielded only once are now retained however. Claassen's initial concern – that item-country parameters would not otherwise be identified – seems to have been too cautious. The resulting dataset includes 4,445 nationally aggregated survey responses gathered by 16 survey projects in 141 countries.

5.2. Other variables

Democracy is measured using the liberal democracy index from the Varieties of Democracy (V-Dem). We use both the point estimates and standard deviations. Analyses of the original dataset use version 8 of the V-Dem dataset; analyses of the extended dataset use version 11 (there is some variation in national liberal democracy scores in various versions of the V-Dem dataset). Several other covariates are included in our structural models:

1. Regional democracy measured using V-Dem's liberal democracy index aggregated to United Nations sub-regions.
2. The logarithm of GDP per capita, included in lagged levels and first differences (the latter being a measure of economic growth). To remove missing values, we use GDP measures from several sources: IMF, World Bank World Development Indicators, Penn World Tables, and Maddison. These data are combined using hierarchical linear models with country-varying intercepts and slopes.
3. Natural resource dependence: an indicator taking a value of one if natural resource products (natural gas, oil, and coal) were valued at more than USD 1,000 per person in any given country-year observation (see, e.g., Haber and Menaldo 2011). Data are primarily drawn from the World Bank World Development Indicators. Missing values are imputed using data from Haber and Menaldo (2011); remaining missing values are given country modes (i.e., 0 or 1).
4. Proportion of the population identifying as Muslim in 1990, from Pew Research.

6. Irregularities in Tai et al "expanded" dataset

A brief consideration of Tai et al's expanded dataset (`exp_claassen_input.rda`, available here: `https://doi.org/10.7910/DVN/XAUF3H`) reveals a number of errors, as well as inconsistencies with Claassen's initial coding rules.

1. Many of the sample sizes appear to be incorrect. The measurement model treats observed survey data as a binomial count, which requires the total sample size and the number of respondents who offered a supportive view of democracy. Many of the sample sizes reported in Tai et al's expanded dataset appear to be incorrect, however, which influences the measurement error attached to each data point in the latent variable model. Some of these errors are obvious, e.g., the implausibly small sample sizes reported in many of the Gallup Voices surveys. For example, a sample of 26 respondents is reported for Iceland in 2004 although the codebook states that the sample size is 502 (see `https://www.icpsr.umich.edu/web/ICPSR/studies/24681/datadocumentation`).

2. Non-representative samples are included. Tai et al's expanded dataset includes samples such as Senegal's 2002 Pew Global Attitudes survey, which Pew themselves describe as not representative of the adult population, but rather, as a "disproportionately urban" sample (see `https://www.pewresearch.org/wp-content/uploads/sites/4/legacy-pdf/165.pdf`). Given that Senegal's population was roughly 60% rural-dwelling in 2002, these data may be considerably biased and should not be included. Many of the other samples fielded this wave of the Pew Global Attitudes surveys were also unrepresentative (e.g., Pakistan, Cote d'Ivoire, India). Other unrepresentative samples, e.g., Morocco and Pakistan, were fielded in the next (2005) wave (see `https://www.pewresearch.org/wp-content/uploads/sites/2/pdf/251.pdf`).

3. Respondents who do not provide a response are (apparently) excluded. Respondents who answered "don't know" or refused to respond appear to have been removed from the Tai et al. expanded dataset before the percentage supporting democracy was calculated. Whether or not

this should be done is perhaps a matter of debate. Dropping such respondents is not, however, consistent with the original Claassen coding scheme. As described in the supplementary materials to Claassen (2020*b*), "all other possible responses (i.e., the difference between the sample size and the number of supportive respondents) were treated, similarly, as not supportive of democracy. These non-supportive respondents may have actively opposed democracy, (e.g., 'an authoritarian government can be preferable to a democratic one'), chosen an intermediate response (e.g., 'for someone like me, it does not matter what kind of government we have'), responded with 'don't know,' or refused to provide any response." (p. 6).

Consider the case of Brazil, the question on whether democracy is important, and the most recent wave of the World Values Survey (2018). The number of observations in the dataset (`WVS_Cross-National_Wave_7_spss_v20200720.sav`) is 1,762. The Tai et al expanded dataset has this sample size as 3,214 however. The raw number of respondents who offered support for democracy in response to this question (defined as offering an opinion above the median on the 1-10 scale; i.e., 6 or above) is 1,271, or 72.1%. Weighting each respondent by the including survey weight, `W_WEIGHT`, produces a weighted response percentage of 72.5%. However the Tai et al dataset reports that 2,557 respondents supported democracy in this question (out of 3,214), for a response percentage of 79.6%. similar results to Tai et al can be obtained by dropping respondents who declined to provide a response to this particular question: excluding these 161 respondents results in an unweighted percentage agreement of 79.4% and a weighted percentage agreement of 79.8%.

7. Further Developing the Unified Model: Issues for Future Research

We recommend the unified model for situations marked by (1) a dynamic cross-national (i.e., time-series, cross-sectional) research design, with (2) one or more latent variables to be estimated using observed indicators. Analyses of time-series, cross-sectional data (TSCS; also known as "panel" data) usually employ features which take advantages of the unique nature of that data, such as within-unit analyses and panel-robust parameter variances. These features are not straightforward to incorporate into the unified model, which is why they are not included in the paper. We discuss these difficulties (and suggest some solutions) below.

7.1. Within-unit analyses

Within-unit analyses are a popular modeling feature for TSCS data. With data such as ours, which varies across countries and years, analysts might estimate structural parameters using only the within-country variance in dependent and independent variables. This helps tackle the problem of country-varying but time-invariant confounds such as historical events (e.g., Claassen 2020*a*). Such an analysis can be carried out by including country fixed effects or by demeaning all variables prior to analysis.

These methods can be applied to Bayesian TSCS models, including the unified model we develop in the present paper. However, by removing between-country (i.e., country-mean) levels of all variables, the data-generating process is fundamentally altered. Yet the function of the unified model is not only to estimate the parameters of a structural model, where within-country analyses might be helpful, but also to estimate one or more of the variables themselves. The unified model is a "generative" model in other words, since it estimates (or "generates") realizations of the variables themselves. As such, a within-country approach appears incompatible with need to estimate latent variables.

A solution perhaps lies in models which simultaneously estimate distinct between-country and within-country structural parameters. One example is the hybrid or within-between model (Bell and Jones 2015; Mundlak 1978). Within a hierarchical framework, this model estimates both within-country parameters for all covariates (including lags of the dependent variable). It also estimates between-country parameters for the same. In principle, it appears that this modeling approach would permit accurate generative estimates of latent variables (by including both within and between-country variance) but also allow a focus on the within-country structural parameters which are of most interest when engaging in causal inference. Incorporating this model within the unified model is not straightforward but looks to be a promising avenue for developing the unified model and rendering it more widely applicable.

7.2. Panel-robust parameter variances

Standard errors produced by classical linear models may be underestimated, e.g., because of heteroskedasticity. Robust methods of estimating standard errors (and variance-covariance matrices more generally) are therefore widely used. A rich variety of similar methods exist for the special case of time-series, cross-sectional data. Here, concerns focus on the clustered nature of the data, where observations are gathered across time in certain countries. The clustering inherent in the data-generating process gives rise to concerns that model residuals may be affected by serial correlation or heteroskedasticity within countries. Panel-robust parameter variance adjustments such as "panel-corrected" standard errors (Beck and Katz 1995) provide robust standard errors while retaining the original parameter estimates.

If within-country analyses were complex to integrate within Bayesian generative models, then robust parameter variances are downright perplexing. One of the few papers on this topic describes it as an "unnatural act" (Startz 2012, 4), "since the posterior arises with no separation of point estimate and distribution around the estimate." There are no variance-covariance matrices to adjust in Bayesian estimation. An alternative approach to this problem is suggested by King and Roberts (2015), who recommend using robust standard errors as a diagnostic tool to identify issues with one's model. They invite researchers to tackle these issues directly via model specification. This approach works well within a Bayesian generative modeling framework, which is not only flexible, but should provide better latent variable estimates to the extent that "missing" model features are incorporated.

At least one of the concerns motivating panel-robust standard errors can be addressed in this way (and is addressed in the particular unified model developed in the present paper), i.e., the serial correlation of residuals. Serially correlated residuals indicate that the dynamics of the dependent variable have not been adequately modeled. This can be addressed by fleshing out the dynamic features which are included, e.g., adding lags of the dependent variable as well as perhaps certain independent variables. In the case at hand, two lags of the dependent variable are motivated by Claassen (2020*a*) in terms of causal inference, but it is noted that they also have the advantage or removing most of the serial correlation in model residuals. Whatever the solution, further work on aligning methods of panel-robust inference from the frequentist econometric tradition and Bayesian generative models for panel data would be of great use to applied researchers.

8. Tables of Results

Table S1. Replication of Claassen (2020*a*) AJPS results

	Dependent variable: Democracy					
	MEX		MOC		UM	
Intercept	−.008	−.003	−.100	−.098	.002	.005
	(.027)	(.027)	(.067)	(.066)	(.003)	(.003)
Democracy$_{t-1}$	1.142	1.144	.582	.583	1.445	1.439
	(.020)	(.020)	(.026)	(.026)	(.038)	(.038)
Democracy$_{t-2}$	−.164	−.165	.352	.352	−.458	−.454
	(.020)	(.020)	(.027)	(.027)	(.038)	(.037)
Mood$_{t-1}$	.009		.015		.010	
	(.003)		(.007)		(.002)	
Mood, democracies only$_{t-1}$		.011		.014		.005
		(.003)		(.007)		(.003)
Mood, autocracies only$_{t-1}$		.003		.019		.018
		(.005)		(.013)		(.004)
Log GDP per capita$_{t-1}$	.001	.000	.010	.010	−.001	.000
	(.003)	(.003)	(.007)	(.007)	(.002)	(.002)
Δ log GDP capita	.152	.151	.141	.149	.097	.093
	(.048)	(.048)	(.138)	(.134)	(.055)	(.053)
Regional democracy$_{t-1}$	.007	.006	.028	.028	.004	.005
	(.004)	(.004)	(.010)	(.009)	(.002)	(.002)
Resource dependence$_{t-1}$	−.011	−.012	−.037	−.035	−.006	−.006
	(.010)	(.010)	(.025)	(.024)	(.006)	(.006)
Proportion Muslim	−.006	−.008	−.021	−.020	−.007	−.005
	(.009)	(.009)	(.022)	(.022)	(.005)	(.006)
N	2435	2435	2435	2435	2300	2300

Original, 1988-2017 dataset used. MEX: measurement error excluded; cell entries are coefficient estimates and conventional standard errors (to facilitate comparison with UM). MOC: method of composition; cell entries are MOC parameter estimates and standard errors (see section 2 for details). UM: unified model; cell entries are posterior means and standard deviations for each parameter. Note that the latent variables variances in the UM differ slightly from presented in the main paper; there, reported parameter estimates for key structural parameters are standardized based on the empirical distributions of latent variables.

Table S2. Replication of Claassen (2020*b*) APSR results

	Dependent variable: Democratic mood		
	MEX	MOC	UM
Intercept	−.037	−.014	−.081
	(.025)	(.083)	(.025)
Mood_{t-1}	1.473	.961	.422
	(.018)	(.030)	(.057)
Mood_{t-2}	−.487	.004	.471
	(.018)	(.029)	(.059)
Democracy_{t-1}	.007	.019	.073
	(.003)	(.010)	(.020)
Δ democracy	−.063	−.003	−.950
	(.021)	(.030)	(.299)
Log GDP per capita_{t-1}	.003	.000	.010
	(.003)	(.009)	(.021)
Δ log GDP capita	.071	.116	1.151
	(.051)	(.173)	(.451)
N	2300	2300	2300

Original, 1988-2017 dataset used. MEX: measurement error excluded; cell entries are coefficient estimates and conventional standard errors (to facilitate comparison with UM). MOC: method of composition; cell entries are MOC parameter estimates and standard errors (see section 2 for details). UM: unified model; cell entries are posterior means and standard deviations for each parameter. Note that the latent variables variances in the UM differ slightly from presented in the main paper; there, reported parameter estimates for key structural parameters are standardized based on the empirical distributions of latent variables.

Table S3. Extension of Claassen (2020*a*) AJPS results

	Dependent variable: Democracy					
	MEX		MOC		UM	
Intercept	−.038	−.031	−.183	−.182	.005	.004
	(.023)	(.024)	(.059)	(.059)	(.002)	(.002)
Democracy$_{t-1}$	1.156	1.157	.641	.641	1.465	1.505
	(.015)	(.015)	(.023)	(.024)	(.032)	(.036)
Democracy$_{t-2}$	−.182	−.182	.283	.284	−.478	−.516
	(.015)	(.015)	(.023)	(.024)	(.032)	(.036)
Mood$_{t-1}$	.008		.015		.007	
	(.003)		(.006)		(.002)	
Mood, democracies only$_{t-1}$		.010		.013		.007
		(.003)		(.007)		(.002)
Mood, autocracies only$_{t-1}$		.001		.017		.005
		(.005)		(.012)		(.003)
Log GDP per capita$_{t-1}$	.006	.005	.019	.019	.002	.002
	(.003)	(.003)	(.006)	(.007)	(.002)	(.002)
Δ log GDP capita	.042	.041	.010	.013	.050	.048
	(.023)	(.023)	(.059)	(.063)	(.021)	(.021)
Regional democracy$_{t-1}$	.007	.006	.029	.028	.004	.003
	(.004)	(.004)	(.009)	(.009)	(.002)	(.002)
Resource dependence$_{t-1}$	−.024	−.025	−.062	−.061	−.012	−.011
	(.008)	(.008)	(.021)	(.021)	(.005)	(.004)
Proportion Muslim	−.010	−.013	−.035	−.035	−.004	−.005
	(.008)	(.008)	(.020)	(.021)	(.004)	(.004)
N	2927	2927	2927	2927	2786	2786

Extended, 1988-2020 dataset used. MEX: measurement error excluded; cell entries are coefficient estimates and conventional standard errors (to facilitate comparison with UM). MOC: method of composition; cell entries are MOC parameter estimates and standard errors (see section 2 for details). UM: unified model; cell entries are posterior means and standard deviations for each parameter. Note that the latent variables variances in the UM differ slightly from presented in the main paper; there, reported parameter estimates for key structural parameters are standardized based on the empirical distributions of latent variables.

Table S4. Extension of Claassen (2020*b*) APSR results

	Dependent variable: Democratic mood		
	MEX	MOC	UM
Intercept	−.035	−.014	−.109
	(.019)	(.083)	(.018)
Mood_{t-1}	1.513	.961	.466
	(.016)	(.030)	(.056)
Mood_{t-2}	−.523	.004	.452
	(.016)	(.029)	(.058)
Democracy_{t-1}	.009	.019	.067
	(.003)	(.010)	(.016)
Δ democracy	−.046	−.003	−.671
	(.017)	(.030)	(.248)
Log GDP per capita_{t-1}	.002	.000	.030
	(.002)	(.009)	(.016)
Δ log GDP capita	.041	.116	.544
	(.019)	(.173)	(.181)
N	2786	2786	2786

Extended, 1988-2020 dataset used. MEX: measurement error excluded; cell entries are coefficient estimates and conventional standard errors (to facilitate comparison with UM). MOC: method of composition; cell entries are MOC parameter estimates and standard errors (see section 2 for details). UM: unified model; cell entries are posterior means and standard deviations for each parameter. Note that the latent variables variances in the UM differ slightly from presented in the main paper; there, reported parameter estimates for key structural parameters are standardized based on the empirical distributions of latent variables.

References

Beck, Nathaniel, and Jonathan N. Katz. 1995. "What To Do (And Not To Do) With Time-Series Cross-Section Data." *American Political Science Review* 89(3): 634–647.

Bell, Andrew, and Kelvyn Jones. 2015. "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross- Sectional and Panel Data." *Political Science Research and Methods* 3(1): 133–153.

Bratton, Michael, Robert Mattes, and E. Gyimah-Boadi. 2005. *Public Opinion, Democracy, and Market Reform in Africa*. Cambridge, UK: Cambridge University Press.

Canache, Damarys, Jeffery J. Mondak, and Mitchell A. Seligson. 2001. "Meaning and Measurement in Cross-National Research on Satisfaction with Democracy." *Public Opinion Quarterly* 65(4): 506–528.

Carpenter, Bob, Andrew Gelman, Matthew Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76(1): 1–32.

Claassen, Christopher. 2019. "Estimating Smooth Country-Year Panels of Public Opinion." *Political Analysis* 27(1): 1–20.

Claassen, Christopher. 2020*a*. "Does Public Support Help Democracy Survive?" *American Journal of Political Science* 64(1): 118–134.

Claassen, Christopher. 2020*b*. "In the Mood for Democracy? Democratic Support as Thermostatic Opinion." *American Political Science Review* 114(1): 36–53.

Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2014. *Bayesian Data Analysis, 3rd Ed.* CRC Press.

Haber, Stephen, and Victor A. Menaldo. 2011. "Do Natural Resources Fuel Authoritarianism? A Reappraisal of the Resource Curse." *American Political Science Review* 105(1): 1–26.

Kellstedt, Paul, Gregory E. McAvoy, and James A. Stimson. 1993/94. "Dynamic Analysis with Latent Constructs." *Political Analysis* 5: 113–50.

King, Gary, and Margaret E. Roberts. 2015. "How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What to Do About It." *Political Analysis* 23(2): 159–79.

Linde, Jonas, and Joakim Ekman. 2003. "Satisfaction With Democracy: A Note on a Frequently Used Indicator in Comparative Politics." *European Journal of Political Research* 42(3): 391–408.

Linz, Juan J., and Alfred Stepan. 1996. *Problems of Democratic Transition and Consolidation: Southern Europe, South America, and Post-Communist Europe*. Baltimore: Johns Hopkins University Press.

Magalhães, Pedro C. 2014. "Government Effectiveness and Support for Democracy." *European Journal of Political Research* 53(1): 77–97.

Mundlak, Yair. 1978. "Pooling of Time-series and Cross-section Data." *Econometrica* 46(1): 69–85.

Norris, Pippa. 2011. *Democratic Deficit: Critical Citizens Revisited*. Cambridge: Cambridge University Press.

Skrondal, Anders, and Sophia Rabe-Hesketh. 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Boca Raton, FL: CRC Press.

Startz, Richard. 2012. "Bayesian Heteroskedasticity-Robust Standard Errors." Working paper, `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.412.3017&rep=rep1&type=pdf`.

Tai, Yuehong 'Cassandra', Yue Hu, and Frederick Solt. 2022. "Democracy, Public Support, and Measurement Uncertainty." *American Political Science Review*. Forthcoming.

Tanner, Martin A. 1996. *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. New York: Springer.

Treier, Shawn, and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52(1): 201–217.