Including Measurement Uncertainty in Time-Series, Cross-Sectional Analyses: The Case of Mood and Democracy

Christopher Claassen*

University of Glasgow

November 2022

*christopher.claassen@glasgow.ac.uk. I would like to thank Sascha Göbel, Armin Seimel and the seminar participants at the University of Bergen and Humboldt University Berlin for helpful comments.

Including Measurement Uncertainty in Time-Series, Cross-Sectional Analyses: The Case of Mood and Democracy

Abstract

Long a feature of microlevel political behavior, the estimation of dynamic, macro-level latent variables is blossoming in comparative politics. Unresolved, however, is the issue of how to incorporate the uncertainty of measurement into subsequent analyses. One proposal is to use the "method of composition," in which multiple samples from the posterior distribution of the latent variable are analyzed, rather than a single point estimate. Using Monte Carlo studies, this paper shows, however, that the method of composition performs poorly in a dynamic context when the latent variable is endogenous to other covariates. In such circumstances (which are likely to be widespread) a unified model that measures latent variables and estimates structural connections in a single step is more accurate and is recommended.

Keywords: measurement error, latent variables, democratic mood, Bayesian statistics

Words: 3,985

1. Introduction

The measurement of macro-level latent variables has seen great advances in recent years. New measures of democracy (Lindberg et al. 2014; Pemstein, Meserve, and Melton 2010), human rights (Schnakenberg and Fariss 2014), and public opinion (Caughey, O'Grady, and Warshaw 2019; Claassen 2019) have generated new insights and rekindled old debates. Moving from static, cross-national measures to dynamic, times-series, cross-sectional measures has allowed scholars to more accurately describe trends and more precisely investigate the causes and consequences thereof. Moving from single-indicator measures to multi-indicator scales has improved measurement accuracy and allowed scholars to quantify the uncertainty of measurement.

It is this latter issue which is now under dispute, and the topic of the present paper. I discuss here the example of democratic mood, i.e., nationally-aggregated public support for democracy. Two recent articles have produced evidence for several links between mood and democracy (Claassen 2020*a*;*b*) but Tai, Hu, and Solt (2022) argue that evidence for these links disappears when measurement uncertainty in mood and democracy is factored in. Scholars have indeed long argued that efforts should be made to factor measurement uncertainty into analyses which use latent variables (see, e.g., Blackwell, Honaker, and King 2017; Fuller 1987). There is little investigation of how this should be done however, particularly in the context of the time-series, cross-sectional designs which feature in analyses of democracy-mood linkages.

The most widespread approach to measurement error in dynamic, cross-national latent variables is to ignore it (e.g., Claassen 2020*b*;*a*), but this likely produces inaccurate parameter estimates to the extent that measurement uncertainty is present. Several studies have therefore turned to a relatively simple method of including measurement uncertainty, the "method of composition" (MOC), which involves running analyses on multiple draws from latent variable posterior distributions rather than utilizing only the point estimates of these distributions (Tai, Hu, and Solt 2022; Treier and Jackman 2008).

I demonstrate here, however, that the method of composition mis-specifies the data generating processes of dynamic latent variables when these are endogenous to other covariates. Instead, I show that it is preferable to measure latent variables and estimate structural linkages with other covariates simultaneously, in a "unified model" (e.g., Fox and Glas 2003; Kellstedt, McAvoy, and Stimson 1993/94). In time-series, cross-sectional situations when latent variables are endogenous to other covariates, Monte Carlo tests reveal that unified models produce more accurate parameter estimates with better uncertainty coverage than the method of composition. Applying this unified model to the case of mood and democracy, I find quite different results from the null findings reported by Tai, Hu, and Solt (2022).

2. Three Approaches to Measurement Uncertainty in Latent Variables

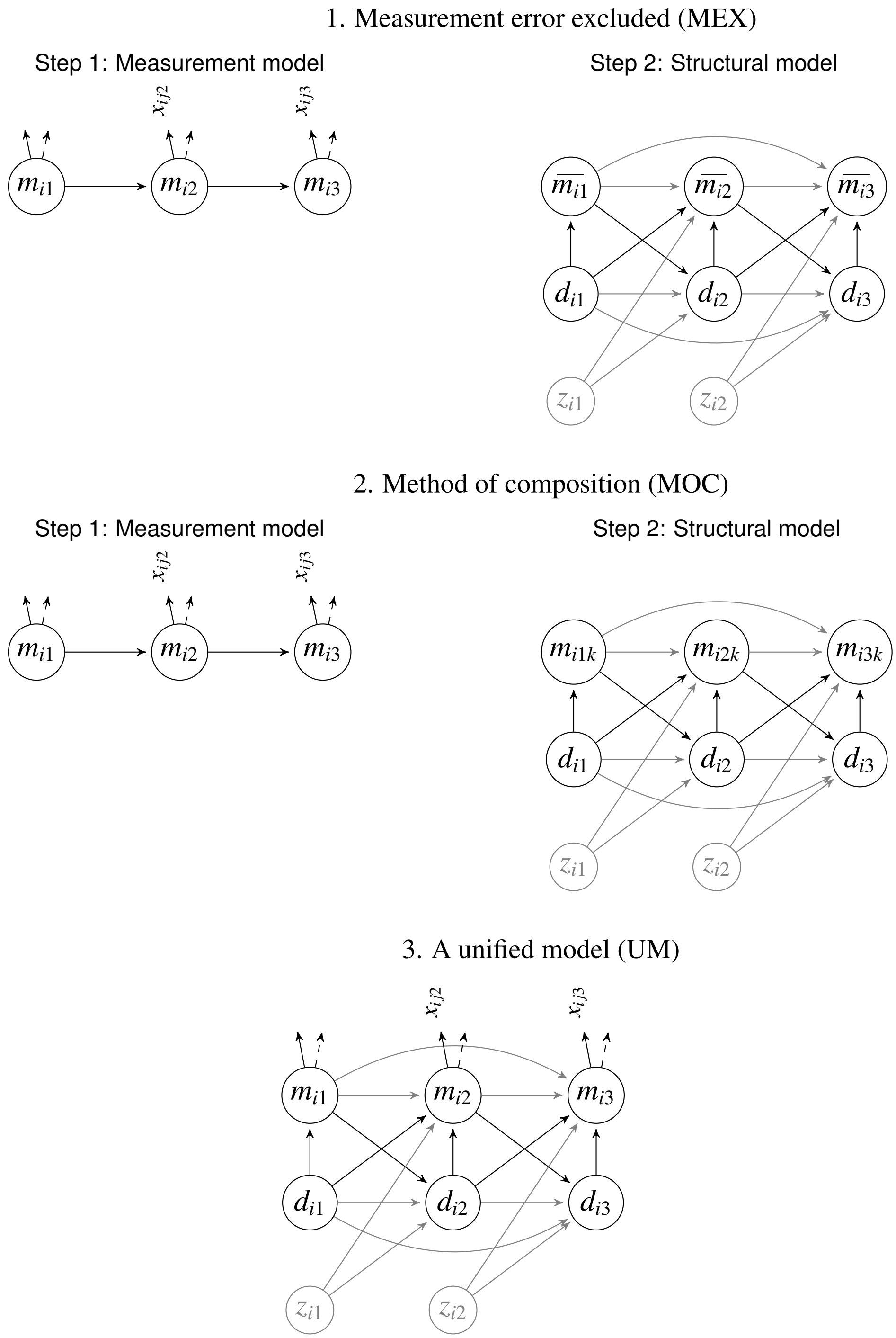
Methods of measuring dynamic latent public opinion have blossomed in recent years (see Caughey and Warshaw 2015; Claassen 2019; McGann 2014). As several analysts have noted, these models allow analysts to produce not only point estimates but also capture the inevitable uncertainty of the measurement process. This measurement uncertainty can then be incorporated into subsequent analyses, allowing for more accurate inferences regarding these latent variables.

However there has been less attention in the discipline to addressing the question of how to include latent variable measurement uncertainty in dynamic macro-level analyses. I outline three approaches that might be taken: measurement error excluded (MEX), the method of composition (MOC), and a unified model (UM). Figure 1 shows how these three methods relate to one another, depicting the time-series, cross-sectional nature of the example at hand using DAGs featuring three time periods.

Simplest of all is to exclude measurement uncertainty, as depicted in panel 1 of Figure 1 and implemented by Claassen (2020*a*;*b*). As can be seen, this requires two separate steps. In the first step, a measurement model is used to produce country by year estimates of mood from aggregated survey data.[1] These data x_{ijt} vary across countries i, survey items j, and years t. There

[1]This measurement model is presented in a greatly simplified fashion here. In practice, the

Figure 1. Three approaches to measurement error in dynamic latent variables



model employed by Claassen (2020*a*;*b*) also includes item parameters and other features, the details of which are described in the supplementary materials.

may be more than one item per country-year, or there may be none at all. Democratic mood in each country and year m_{it} is treated as the latent variable underlying the observed survey indicators in each country-year unit. As the figure indicates, latent mood is also assumed to follow a type of first-order autoregressive process, specifically a random walk.

When measurement error is excluded (MEX), analysts simply extract point estimates of the latent variable; in this case, estimates of mood for each country and year. The measurement model in this example is estimated using Markov Chain Monte Carlo methods, in which a high-dimensional probability distribution is sampled. Some of all of these MCMC samples are saved. Point estimates are obtained by, e.g., finding the mean of the k MCMC samples for each country and year in which latent estimates are required. This is shown in step 2 of the first approach. These point estimates are included within a structural regression model.

In the example at hand, the structural model of democracy is taken from Claassen (2020*a*) and the model of mood is taken from Claassen (2020*b*); they are combined here as summarized as in the figure. Mood is assumed to be a function of its previous two lags, both lagged and immediate effects of democracy d_{it}, as well as the effects of some lagged covariates, z_{it-1}. The lagged effect of democracy on mood is described by Claassen (2020*b*) as the *socialization* effect as it accumulates over time in this dynamic set-up. The immediate effect of democracy (specified using a first difference in Claassen (2020*b*)) is *thermostatic effect* of change in democracy.

Democracy is also a dependent variable in this structural model: it is modeled as endogenous to its previous two lags, lagged mood, plus some lagged covariates. The effect of lagged mood on democracy is labeled as the *Lipset* effect given Lipset's early discussion of the stabilizing role played by popular democratic support (e.g., Claassen 2020*a*). Note that democracy may also be measured with uncertainty, but this is not depicted in Figure 1.

Next I consider a method of including measurement uncertainty, described initially by Tanner (1996) as the "method of composition" (MOC), introduced to political science by Treier and Jackman (2008), and employed by Tai, Hu, and Solt (2022). This method is presented in the second panel of Figure 1. Two steps are still used. And the first step is identical to the corresponding step in the first approach: a measurement model is applied to observed indicators to estimate the latent variable. However, when preparing for the second step, the analyst now saves a number of samples k from the MCMC sampling of the posterior distribution of the latent variable. In the second step, structural models are run k times, each time using one of the samples drawn from the latent variable. Results from these k models are then combined to reflect both parameter and measurement uncertainty.[2]

Yet a third approach exists, estimating measurement and structural models in one step. Such a unified model is hardly novel: attempts to factor measurement uncertainty into subsequent analysis go back to Jöreskog's LISREL model in the 1970s and the subsequent emergence of what became known as "Structural Equation Modeling" (SEM; Skrondal and Rabe-Hesketh 2004). More complex variations of this basic idea have been proposed, e.g., item-response theoretic measurement models combined with hierarchical structural models (e.g., Fox and Glas 2003) or dynamic latent variables with multiple exogenous determinants (i.e., the DYMIMIC model of Kellstedt, McAvoy, and Stimson 1993/94). The essence of these approaches is to construct a single model in which latent variables are measured and their structural relationships (i.e., the linkages between the latent variable and other covariates) are estimated (see panel three of Figure 1). As Figure 1 shows, the unified model (UM) combines the two steps of measurement and structural estimation into one. These are estimated simultaneously, as part of a single likelihood.

The advantage that UM holds over two-step methods such as MOC (and MEX) can be seen in Figure 1. Two-step methods require two distinct data-generating processes for the latent variable to be assumed: typically a simpler process ignoring other covariates when it comes to measurement (step 1), and a more complex process potentially involving exogenous covariates when it comes to estimating structural parameters (step 2). Both DGPs cannot be correct. Indeed, if other variables (e.g., d_{it}) are believed to exert an influence on the latent variable (e.g., m_{it}), then these should be included in the measurement model.[3] This is exactly the logic of the unified model (and its

[2]See the supplementary materials for a more detailed description of this method.

[3]Blackwell, Honaker, and King (2017) make this same point in motivating their method of

conceptual cousins, SEM, DYMIMIC, etc.) and the rationale for why it offers a superior method of estimating structural linkages between a latent variable and other covariates while incorporating measurement error.

3. Monte Carlo Study

We turn now to a test of the accuracy the three approaches. A Monte Carlo approach is used, where a data-generating process capturing the main features of the case of interest is set up, and the accuracy of the three methods of including measurement error in retrieving some parameters of interest are tested (see, e.g., Carsey and Harden 2014). The DGP involves a latent variable varying across cross-sectional ($N = 50$) and temporal ($T = 30$) units, observed only via a fragmented and partial set of 3 indicators (only 20% of the possible NT observations for each are observed).[4] I consider three distinct scenarios, running a separate Monte Carlo study for each:

Scenario 1. The latent variable is exogenous, with its lagged value exerting a positive effect (denoted as δ_1) on a second, observed variable, but with no reciprocal effect from the observed variable.

Scenario 2. The latent variable is endogenous to the observed variable, with the lagged observed variable exerting a positive effect on the latent variable (δ_2) and the immediate first difference exerting a negative effect (δ_3). The observed variable is exogenous.

Scenario 3. Both latent and observed variables are endogenous, with all three structural parameters (δ_1, δ_2, and δ_3) in play.

---

including measurement uncertainty, "multiple overimputation." I do not include this method here because it does not allow for a customized dynamic latent variable measurement model – such as that presented by Claassen (2019) – to be included.

[4] See the online supplementary materials for further details.

These situations are designed to capture the possible relationships between democracy and democratic mood. The most complex third scenario allows for the three distinct links between democracy and democratic mood which were identified by Claassen (2020*a*;*b*): δ_1 is comparable to the *Lipset effect*, the lagged effect of mood on democracy; δ_2 parallels the *socialization effect* of lagged democracy on mood; and δ_3 is comparable to the *thermostatic effect* of change in democracy on mood. Scenarios 1 and 2 provide simpler data-generating processes in which either of the two variables is exogenous.

In all three scenarios, I treat the latent variable as unobserved. What is observed are three indicators of this variable. Each takes a TSCS form, but each is sparsely distributed across temporal and spatial units. I assume that observations of each indicator exist only for 20% of the 30 temporal and 50 spatial units, with the remaining 80% of possible observations being missing. This approximates the survey measures that are available for support for democracy, which are scattered across time and space in the form of different survey items.

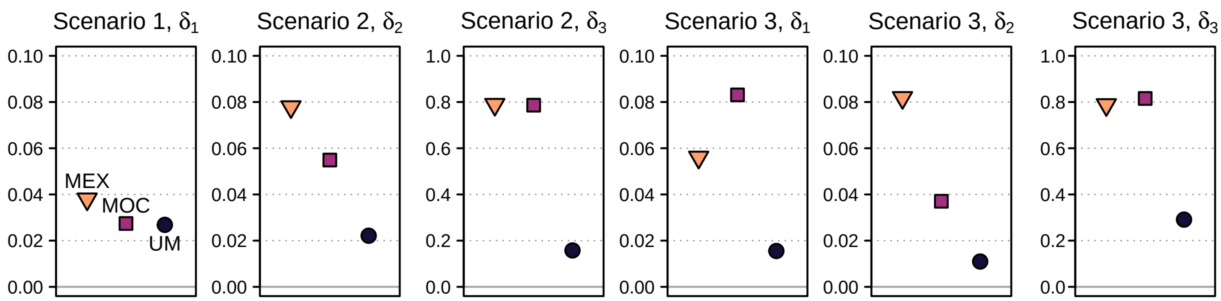
In our Monte Carlo studies, each of the methods (MEX, MOC, and UM) are applied to 500 randomly generated datasets. For each of the three methods and randomly generated datasets, I save 500 draws from the posterior distributions of the structural parameters (δ_1, δ_2, and/or δ_3): this allows for measurement error and structural modeling error to be included. Across our three scenarios, three methods, and three parameters, two quantities of interest are tracked: the bias of each method and its 50% uncertainty (confidence or credible) interval coverage (e.g., Carsey and Harden 2014). These results are presented in Figure 2.

In scenario 1, where the latent variable is exogenous to the observed variable, both MOC and UM perform similarly well: these two estimators are less biased than when measurement error is exclude (MEX; top row of Figure 2). In both cases the uncertainty interval coverage is substantially below the nominal level of 50%; nevertheless, both MOC and UM are preferable to MEX here.

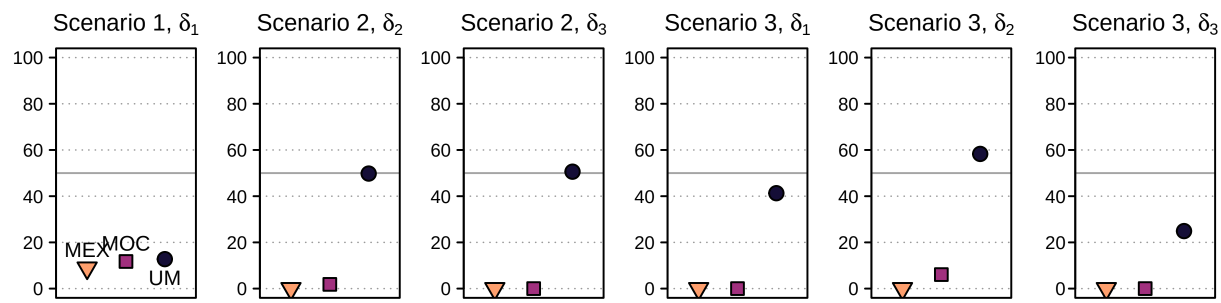
In scenario 2, where the latent variable is treated as endogenous to the observed variable, UM now shows a marked advantages over MOC (and MEX). MOC shows much more bias than

Figure 2. Results of the Monte Carlo studies: bias and uncertainty interval coverage

(a) Bias (Mean Absolute Error) – lower is better



(b) 50% Uncertainty Interval Coverage – closer to 50 is better



Bias and uncertainty interval coverage of the three methods – (1) Measurement Error Excluded (MEX), (2) Method of Composition (MOC), and (3) a Unified Model (UM) – in estimating the three structural parameters (δ_2, δ_1, and δ_3) under three different scenarios. In scenario 1, the latent variable is exogenous to the observed variable; in scenario 2, the former is endogenous; in scenario 3, both are endogenous. Results are averaged across the 500 iterations of each Monte Carlo study.

UM in retrieving the effects of the lagged observed variable on the latent variable, δ_2, and the effect of the first-differenced observed variable on the latent variable, δ_3. MOC (and MEX) show very poor uncertainty interval coverage, especially compared to UM, where coverage is accurate.

Finally, in the most complex scenario, the third, similar results to the first two Monte Carlo studies are observed. For all parameters, UM shows the least bias and the best uncertainty interval coverage.

In sum, our Monte Carlo studies show that when a dynamic latent variable is endogenous (i.e., scenarios 2 and 3) UM is more accurate than MOC at retrieving estimates of structural parameters.[5] When such a dynamic latent variable can be assumed to be exogenous to other covariates

[5]Additional results are reported in the online supplementary materials.

both UM and MOC perform similarly well.

4. Empirical Application: Democracy and Democratic Mood

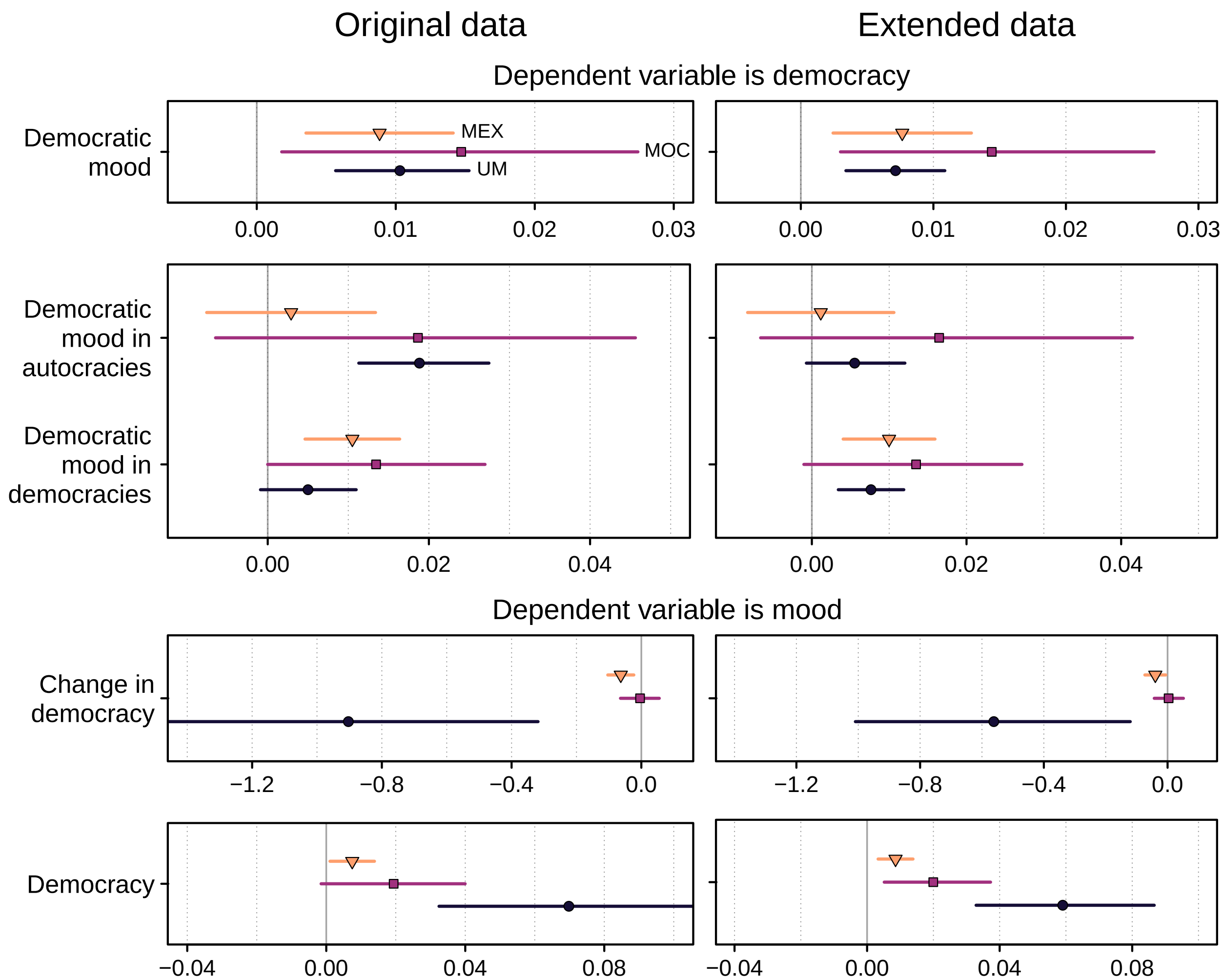
In the final component of this paper, I replicate the models of Claassen (2020*a*) and Claassen (2020*b*) using the three approaches. In doing so, I extend the analyses of Tai, Hu, and Solt (2022) by adding the third method of incorporating measurement error, UM. For the purposes of this comparison of methods, I use a pooled design without country fixed effects. Two datasets are examined: first, the original dataset gathered by Claassen (2019; 2020*a*;*b*) which ranged from 1988 to 2017; second, an extended dataset which adds data from 2018 to 2020 and some additional survey projects.[6]

Figure 3 shows the results of these analyses. When democracy is the outcome (top half of Figure 3), the main link of interest relates to "supportive" effect of mood (i.e., the Lipset effect and parameter δ_1 in the Monte Carlo analysis reported in the previous section). It can seen that these estimates are largely consistent across the three approaches when considering the overall effect of mood (top panel of Figure 3), and consistent also with the pooled estimate reported by Claassen (2020*a*, Model 1.1). The parameter estimates are somewhat less consistent when allowing the effect of mood to vary across regimes (second panel of Figure 3). Indeed, the UM estimates, which I have shown to be the most accurate in our Monte Carlo test, depart from the original estimates reported by Claassen (2020*a*, Model 1.2), which pointed to a positive and significant effect of mood in democracies, but not autocracies (i.e., the opposite pattern observed here). When applied to the extended dataset, UM produces results consistent with Claassen (2020*a*).

When mood is treated as the dependent variable (lower half of Figure 3), there are two parameters of interest: the effect of the lagged level of democracy (i.e., the socialization effect;

[6]Note that Tai, Hu, and Solt (2022) also create an expanded dataset. However I found errors in that dataset as well as inconsistencies with how measures were coded in the original. See the supplementary materials for further details.

Figure 3. Estimated linkages between mood and democracy, three approaches



Estimated linkages between mood and democracy, using three approaches for handling measurement error: (1) Measurement Error Excluded (MEX); (2) Method of Composition (MOC); and (3) a Unified Model (UM). Points show estimated effects of key covariates (in rows) on each dependent variable (panels); horizontal bars show 95% confidence or credible intervals. For the MEX and MOC results N is 2,425 (original data) and 2,927 (extended) when the dependent variable is democracy and 2,300 and 2,786 when the dependent variable is mood. Because the UM estimates both mood and democracy simultaneously, the N is the same for models of both democracy and mood: 2,300 for the original data and 2,786 for the extended data. Each model includes additional covariates and lags of the dependent variable; see the supplementary materials for full results.

parameter δ_2) and the effect of an immediate change in democracy (i.e, the thermostatic effect; parameter δ_3). Estimates of the socialization effect are generally positive here, increasing in magnitude as the results proceed from MEX to MOC to UM (the MOC estimates based on the original dataset are positive although insignificant). This is consistent with the slightly positive and significant effect reported in Claassen (2020*b*, Model 1.1).

Estimates of the thermostatic effect show substantial variation across the three approaches: MEX returns small negative estimates, significant using the original data, insignificant otherwise; MOC produces estimates which are more or less exactly zero; the UM shows strongly negative effects. Since the Monte Carlo tests found UM to be substantially better than MOC and MEX in estimating first-difference effects, the strong negative results obtained here would appear to be the most accurate estimates of the effect of change in democracy on mood. They are also consistent with the results reported by (Claassen 2020*b*).

Substantively, the application of UM to the mood and democracy datasets produces similar results to the pooled results reported by Claassen (2020*a*;*b*): mood has a positive association with subsequent change in democracy and change in democracy (i.e., the Lipset effect) is negatively associated with subsequent change in mood (the thermostatic effect). However, results are less consistent when it comes to the regime-varying linkages between mood and democratic change: while Claassen (2020*a*) reported a positive effect only within democracies, our UM analyses show such an effect only in the extended dataset; in the original dataset the effect is insignificant.

An important caveat is that I have used pooled TSCS designs in these analyses. As Claassen (2020*a*) argues, interpreting associations between mood and democratic change in a causal fashion additionally requires accounting for unobserved time-invariant confounders such as historical junctures affecting institutions and political cultures. Although solutions such as fixed-effect estimators are simple to implement in two-step approaches, they are less straightforward in dynamic generative models, such as those employed in our unified model.[7] This is an important area for future research.

[7] Mood is estimated at the same time as its structural linkages with democracy; if a within-country design is used to estimate the latter, the measurement of mood is also altered.

5. Conclusion

This paper considers the issue of measurement error in dynamic latent variables, focusing on the example of links between democracy and democratic mood. Three approaches to handling measurement error are discussed: excluding measurement error (MEX; Claassen e.g. 2020*a*;*b*), including measurement error via the method of composition (MOC; Tai, Hu, and Solt 2022; Tanner 1996; Treier and Jackman 2008), and including measurement error via a unified model (UM; Fox and Glas 2003; Kellstedt, McAvoy, and Stimson 1993/94). These three approaches are discussed conceptually, tested using Monte Carlo analyses, and applied to the real-world case of democracy and mood.

I agree with Tai, Hu, and Solt (2022) that including measurement error affects results in downstream structural models; it should generally be attempted (unless, for example, such measurement error is negligible). Yet I also identify a significant weakness of their preferred approach, MOC. When latent variables are endogenous to other variables (e.g., mood being endogenous to existing levels of democracy), MOC, like MEX, requires the latent variable be modeled using two different data-generating processes: a simpler DGP ignoring other variables for the purposes of measurement, and a more complex DGP for second-step estimation of structural parameters. The Monte Carlo analyses conducted in the paper show that this Janus-faced quality has deleterious consequences, with MOC performing poorly in comparison with UM, occasionally as poorly as MEX. MOC is therefore not suitable for including measurement error in dynamic latent variables when these are endogenous to other covariates.

Unless researchers can assume that dynamic latent variables are exogenous, it is preferable to use a unified modeling framework to simultaneously estimate the latent variable and its structural linkages with auxiliary covariates. Single step unified models are more complex to run than two-step methods such as MOC, but they appear to be necessary to obtain accurate parameter estimates when using dynamic latent variables measured with error.

References

Blackwell, Matthew, James Honaker, and Gary King. 2017. "A Unified Approach to Measurement Error and Missing Data: Overview and Applications." *Sociological Methods & Research* 46(3): 303–41.

Carsey, Thomas M., and Jeffrey J. Harden. 2014. *Monte Carlo Simulation and Resampling Methods for Social Science*. Sage.

Caughey, Devin, and Christopher Warshaw. 2015. "Dynamic Estimation of Latent Opinion Using a Hierarchical Group-Level IRT Model." *Political Analysis* 23(2): 197–211.

Caughey, Devin, Tom O'Grady, and Christopher Warshaw. 2019. "Policy Ideology in European Mass Publics, 1981–2016." *American Political Science Review* 113(3): 674–693.

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.

Fox, Jean-Paul, and Cees A. W. Glas. 2003. "Bayesian Modeling of Measurement Error in Predictor Variables Using Item Response Theory." *Psychometrika* 68(2): 169–91.

Fuller, Wayne A. 1987. *Measurement Error Models*. New York: Wiley.

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

Lindberg, Staffan I., Michael Coppedge, John Gerring, Jan Teorell, Daniel Pemstein, Eitan Tzelgov, Yi ting Wang, Adam Glynn, David Altman, Michael Bernhard, Steven Fish, Alan Hicken, Matthew Kroenig, Kelly McMann, Pamela Paxton, Megan Reif, Svend-Erik Skaaning, and Jeffrey Staton. 2014. "V-Dem: A New Way to Measure Democracy." *Journal of Democracy* 25(3): 159–169.

McGann, Anthony J. 2014. "Estimating the Political Center from Aggregate Data: An Item Response Theory Alternative to the Stimson Dyad Ratios Algorithm." *Political Analysis* 22(1): 115–129.

Pemstein, Daniel, Stephen Meserve, and James Melton. 2010. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." *Political Analysis* 18(4): 426–49.

Schnakenberg, Keith E., and Christopher J. Fariss. 2014. "Dynamic Patterns of Human Rights Practices." *Political Science Research and Methods* 2(1): 1–31.

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

Tai, Yuehong 'Cassandra', Yue Hu, and Frederick Solt. 2022. "Democracy, Public Support, and Measurement Uncertainty." *American Political Science Review* `https://doi.org/10.1017/S0003055422000429`.

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.