# Effects of Measurement Error on Autoregressive Parameters in Time Series Analysis

UC DAVIS
UNIVERSITY OF CALIFORNIA

Kristine D. Christianson[1], Siwei Liu[2], Emilio Ferrer[1]
[1]Department of Psychology, University of California, Davis
[2]Department of Human Ecology, Human Development, University of California, Davis

## BACKGROUND

In psychology, interest in investigating intraindividual processes through individual time series analysis has grown in the last two decades. However, despite the fact that research has shown that failure to account for measurement error leads to biased estimates of autoregressive parameters (e.g., Schenker & Gentleman, 2001; Shuurman, Houtveen, & Hamaker, 2015), the majority of these analyses do not explicitly incorporate measurement error. This is problematic because the variables typically used in the social sciences are difficult or impossible to measure directly, meaning that measurement error is omnipresent in these fields.

## PURPOSE

The purpose of this project is to investigate the effects of measurement error in analysis of individual time series data by comparing the dynamic factor analysis (DFA) model to autoregressive (AR) and autoregressive + white noise (AR + WN) models using composites of observed indicators. In this project, we are interested in the following research questions:

1. How does unique variance affect dynamic parameters in time series analysis?
2. Does the AR + WN model recover the AR parameter as well as the DFA model?
3. How do these models perform compared to an AR model which does not incorporate measurement error?

## METHOD

Monte Carlo Simulation

- Data were generated from a lag-1 DFA model with 1 latent factor (DFA[1,1]) using the MARSS package in R (R Core Team, 2016)
  - 500 data sets generated per condition
- The DFA(1,1) was fit to the generated data and lag-1 AR (AR[1]) and lag-1 AR + WN (AR[1] + WN) models were fit to composites of the indicators from the DFA(1,1) model
- The following parameters were varied across conditions:
  - *Factor loading/unique variance ratio*
    - .5/.75, .7/.51, and .9/.19
  - *Number of indicator variables*
    - 3, 5, and 7
  - *Magnitude of AR parameter*
    - .2, .5, and .8
  - *Number of time points*
    - 50, 100, and 500

## RESULTS

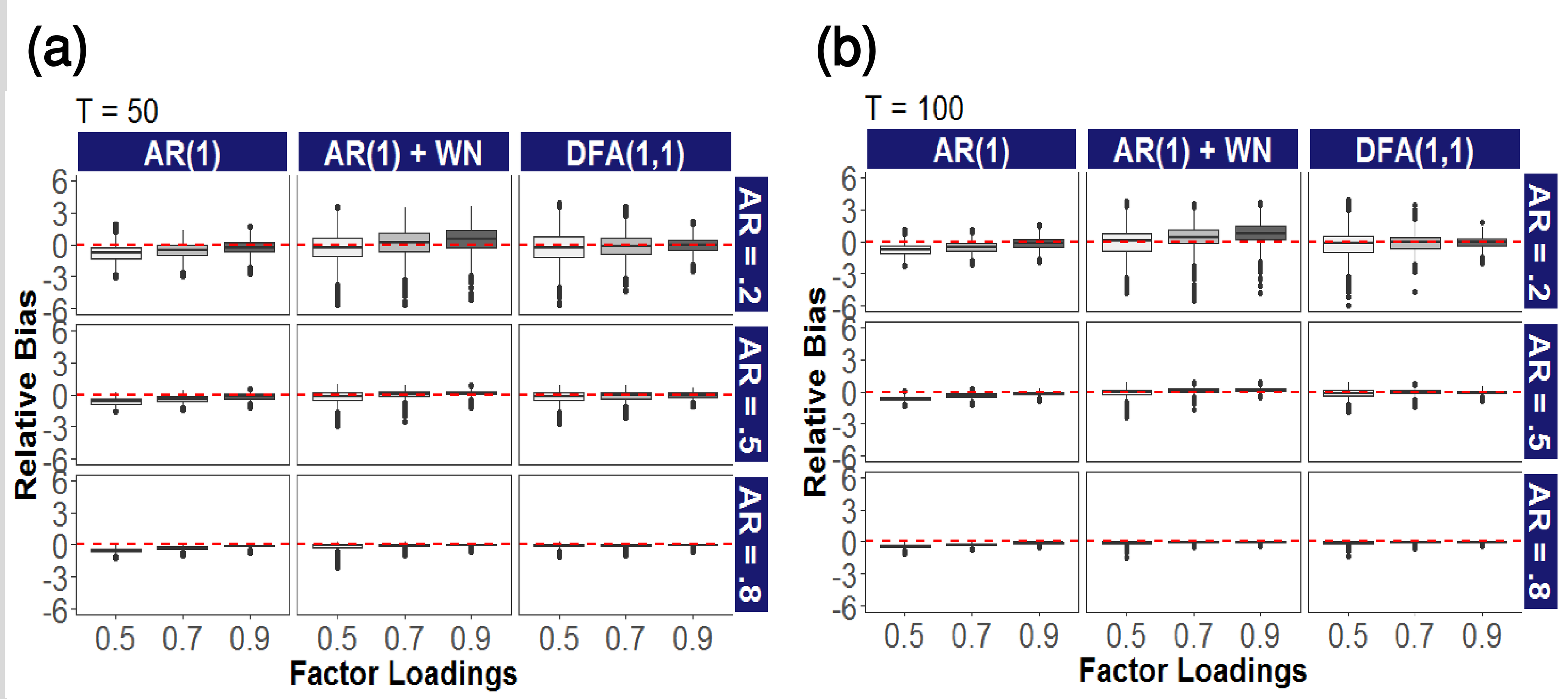
### Evaluating Model Performance

The performance of each of the models was assessed by examining the relative bias and the 95% coverage rates for the AR parameters from the DFA(1,1), AR(1), and AR(1) + WN models. Relative bias was computed as $(\hat{\phi} - \phi)/\phi$, where $\hat{\phi}$ is the estimated value of the AR parameter and $\phi$ is the true parameter. Coverage rates were formulated by examining the proportion of times the true value of the AR parameter occurred within the 95% confidence intervals of the maximum likelihood estimated values in each of the simulated conditions.

### Model Convergence

A number of the models failed to converge across some of the simulation conditions. For the DFA(1,1) and AR(1) + WN models, convergence greatly improved as the number of time points and the AR parameter increased, and as the amount of unique variance decreased. The AR(1) model converged 100% of the time for all of the simulation conditions examined. Results are reported for the converged solutions only.

### Relative Bias



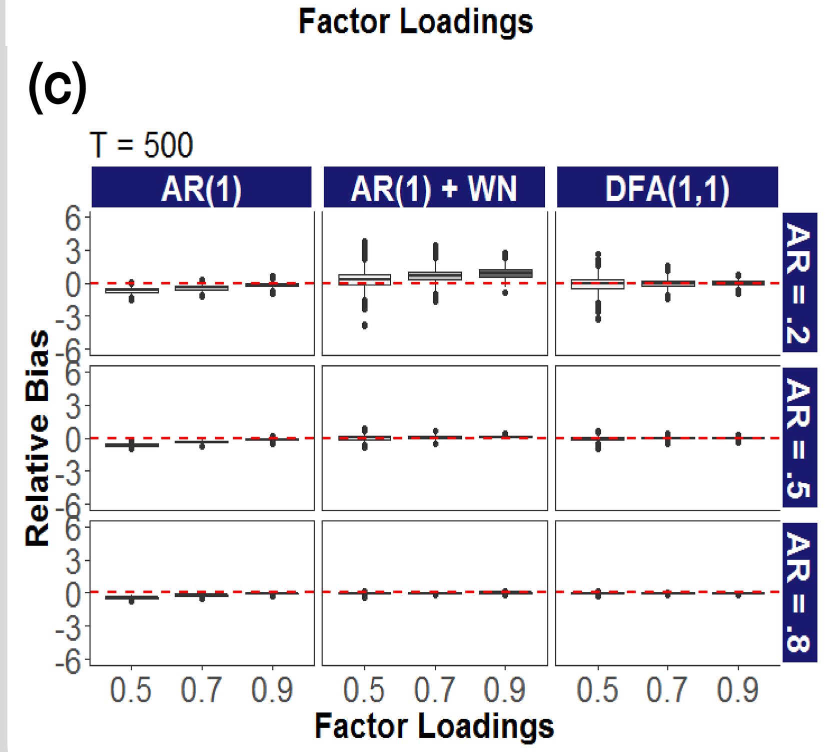




*Figure 1.* Relative bias of AR parameter estimates across the converged solutions for length of time series equal to (a) 50, (b) 100, and (c) 500. There were no notable differences between various numbers of indicators, so results are collapsed across these conditions.

## RESULTS

### Coverage

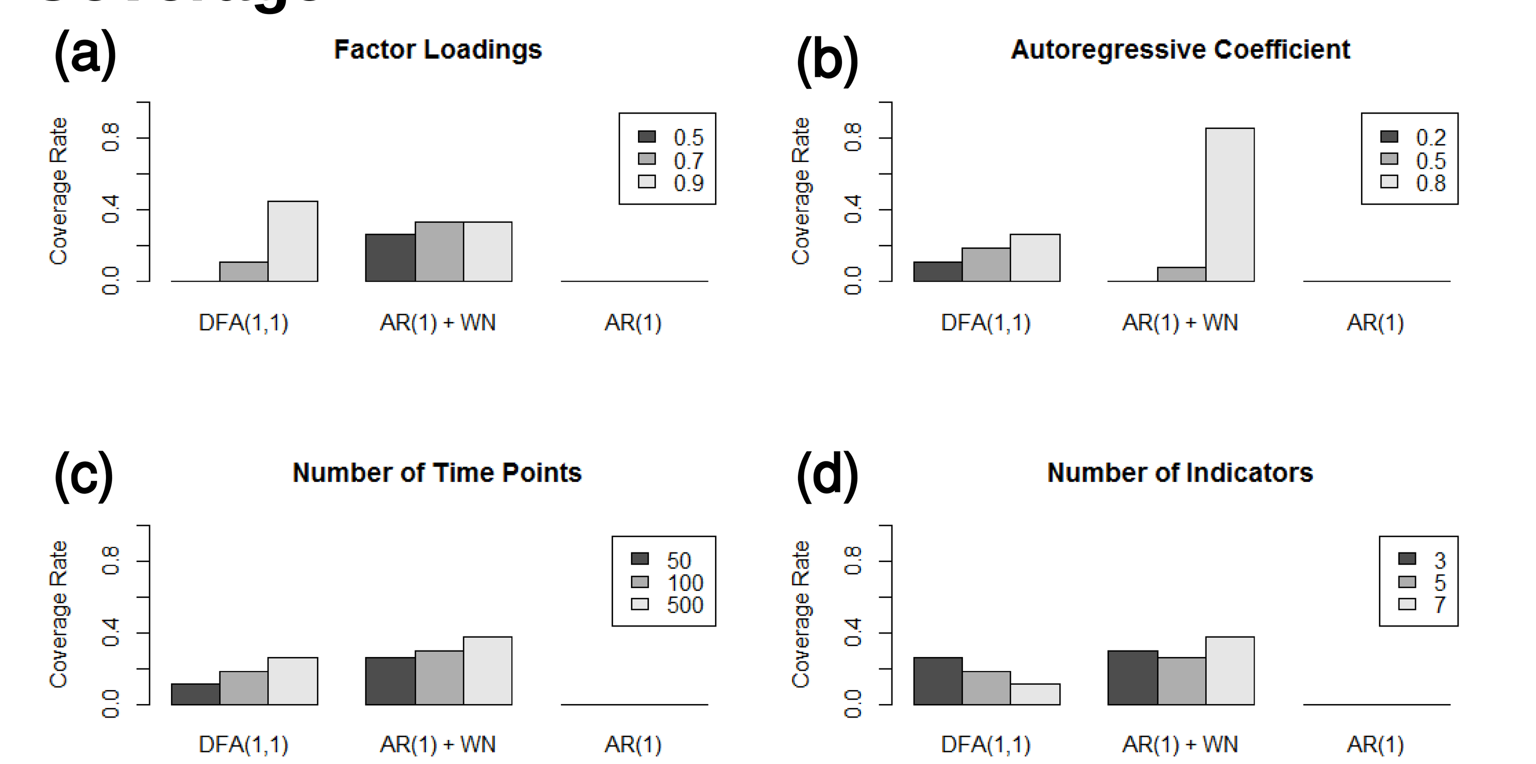



*Figure 2.* 95% coverage rates for DFA(1,1), AR(1) + WN, and AR(1) models with (a) magnitude of factor loadings varied, (b) magnitude of AR parameters varied, (c) number of time points varied, (d) number of indicators varied. Reported coverage rates are collapsed across other conditions.

## DISCUSSION

When the true AR parameter is high and the time series is sufficiently long (e.g., $T > 100$), the AR(1) + WN model performs as well, or nearly as well, as the DFA(1,1) model at recovering the true AR parameter. Both the amount of relative bias and the 95% coverage rates for the estimates of the AR parameter seemed to improve with increasing length of time series and when there was a smaller proportion of unique variance. The AR(1) model tended to severely underestimate the AR parameter in most of the simulation conditions.

## CONCLUSIONS

Overall, these findings suggest that in the absence of multiple indicators, the AR(1) + WN model may be a reasonable model for incorporating measurement error in time series analysis as long as convergence criteria are met. The AR(1) model, on the other hand, tends to lead to attenuation of the AR estimate in the presence of measurement error.

## CONTACT

For more information, or for a copy of this presentation, please contact Kristine Christianson at kdchristianson@ucdavis.edu.