Advance Access publication February 24, 2015

Political Analysis (2015) 23:278–298
doi:10.1093/pan/mpu020

A Continuous-Time, Latent-Variable Model of Time Series Data

Alexander M. Tahk

Department of Political Science, University of Wisconsin–Madison, 1050 Bascom Mall, Madison, WI 53706
email: atahk@wisc.edu

Edited by Patrick Brandt

Many types of time series data in political science, including polling data and events data, exhibit important features—such as irregular spacing, noninstantaneous observation, overlapping observation times, and sampling or other measurement error—that are ignored in most statistical analyses because of model limitations. Ignoring these properties can lead not only to biased coefficients but also to incorrect inference about the direction of causality. This article develops a continuous-time model to overcome these limitations. This new model treats observations as noisy samples collected over an interval of time and can be viewed as a generalization of the vector autoregressive model. Monte Carlo simulations and two empirical examples demonstrate the importance of modeling these features of the data.

1 Introduction

Most time series models assume that time is divided into discrete, regularly spaced units. At each unit of time, the observations are perfect measures of the variables of interest. These assumptions are quite reasonable when the process is directly measured and truly occurs at discrete units of time.

If the data consist entirely of these variables, then there may be little wrong with standard models. However, many types of data do not meet this description and examples abound in political science. Two common examples are polling data to measure public opinion and events data to measure the level of cooperation or conflict between countries. Both types of data are noisy and irregularly spaced, while the concepts they measure are continuous-time processes.

The ability to collect data as totals over regularly spaced intervals does not imply that such an approach is ideal. Even if yearly totals are available for all variables, some variables may be available over more fine-grained time intervals. The use of less fine-grained data may mean the loss of some information about the relationship between variables. It can also introduce other problems if the data are not truly totals of the variable of interest over the time intervals—as when all the polls taken during a month are averaged to create a single monthly observation.

1.1 *Data Problems*

Before discussing models to deal with these features, it may be helpful to define the specific data features that will be discussed:

- Multiple time series**:** The data are multivariate in nature.
- Measurement error**:** Observations are not perfect measures of the variables of interest. They are noisy measures of a latent process which is never directly observed.

Author's note: The author wishes to thank John Ahlquist, Patrick Brandt, Kosuke Imai, Simon Jackman, Jon Krosnick, Suzanna Linn, Jon Pevehouse, James Sieja, Susannah Tahk, David Weimer, and participants at Princeton University's Political Methodology Research Seminar and the 28th Annual Meeting of the Society for Political Methodology. Supplementary materials for this article are available on the *Political Analysis* Web site.

© The Author 2015. Published by Oxford University Press on behalf of the Society for Political Methodology.
All rights reserved. For Permissions, please email: journals.permissions@oup.com

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

- Mixed-frequency observations: Each series is regularly spaced but this spacing differs between series.
- Irregularly spaced observations:[1] The time from one observation to the next is not constant. This can occur for two different reasons. The observations may be regularly spaced with missing observations or the observations may be irregularly spaced without an underlying sampling interval (Jones 1985).
- Noninstantaneous observations**:** An observation measures an aggregate, such as a sum or average, of a latent process over an interval of time rather than the value of that latent process at a single point of time.
- Overlapping observations**:** One noninstantaneous observation is a measurement over a time interval that overlaps the measurement time or time interval of another observation.[2] The second observation may be instantaneous or noninstantaneous.

Some of these features may be ignored when sufficiently small. Ignoring measurement error may have little effect when it is very small in comparison to other sources of variation in the data. Nonetheless, there are many circumstances where these features cannot be safely ignored. Many time series methods deal with data containing one or more of these features—particularly the first two—but not all. Although no statistical model can capture all data features, this article seeks to develop a method that is applicable to data that contain any and all of these six features discussed above.

1.2 *Polling Data*

Opinion polls are widely used in the study of public opinion. However, they have several important features that are ignored or inadequately addressed in most time series studies. First, polls are a noisy measure of public opinion due to sampling error and other factors. Second, polls are often irregularly spaced. Even when regularly spaced, their spacing may not match up with other variables of interest. Third, polls are not instantaneous measures of public opinion. Instead, the interviews may be conducted over days, weeks, or months and may overlap with other polls or measurements.

1.3 *Events Data*

Events data are common in quantitative international relations. These data describe the actions of one international actor toward another. Events are coded from news reports and are commonly translated onto an interval scale.

Events data present many of the features discussed above. Measurement error accrues from errors in news reports, coding errors, the choice of possible event codes, and their translation onto an interval scale. Because of their dependence on news reports, observations in events data are commonly irregularly spaced. A period without a news report does not necessarily indicate that no event occurred during the period. Observations are also often noninstantaneous because they describe the actions taken over a period of time rather than identifying a precise moment for each action.

1.4 *Other Data*

All of these problems extend to many other types of data. Measurement error is almost ubiquitous in social science data, as few variables are viewed as perfect measurements of the desired variable. Gross domestic product (GDP), inflation, and other flow variables, which describe changes over

[1]Common synonyms for "irregularly spaced" in the time series literature include "unequally spaced" and "unevenly spaced."

[2]For example, one observation might be a sum of a latent process from time t_1 to t_3 while another is a sum from time t_2 to t_4 or an instantaneous measurement at time t_2 where $t_1 \leq t_2 \leq t_3$.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

time, have inherently noninstantaneous measurements. Count data also frequently have noninstantaneous measurements. Many of these variables are collected over changing lengths of time or available only for certain periods of time but not others, leading to irregular spacing. If one series is irregularly spaced while others are not, the collective set of series must be regarded as irregularly spaced. Even when individual series are regularly spaced, combining data from many sources often leads to mixed-frequency observations.

2 Existing Methods

This section provides a brief overview of several relevant existing models. Two of these are common in political science: vector autoregression (VAR) and Stimson's dyad ratios algorithm. Some uses of these models can lead to biases from temporal aggregation and imputation, which are also discussed in this section and illustrated in the Monte Carlo studies in Section 4. The two other models discussed in this section are relatively rare in political science but relevant to dealing with some of the data features listed in Section 1.1. These are state-space models and continuous-time models. This article builds on both of these models.

2.1 *Vector Autoregression*

Analyzing multiple time series with VAR has become standard in political science since being introduced to the discipline in work such as Freeman et al. (1989). A VAR model assumes that the values observed at time t are a function of lagged variables from p previous periods of time as well as a vector of errors that are uncorrelated across time (though elements of this vector at a given time may be correlated). A pth order vector autoregressive or VAR(p) model takes the form

$$\mathbf{y}_t = \boldsymbol{\alpha} + \mathbf{B}_1\mathbf{y}_{t-1} + \ldots + \mathbf{B}_p\mathbf{y}_{t-p} + \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0, \mathbf{T}), \tag{1}$$

where $\mathbf{y}_t$ is the $k \times 1$ vector of dependent variables at time t, $\boldsymbol{\alpha}$ is the vector of intercepts, $\mathbf{B}_i$ is the $k \times k$ matrix of coefficients on lags of order i,[3] ε_t are the $k \times 1$ error terms at time t, and $\mathbf{T}$ is the $k \times k$ variance–covariance matrix of the error terms.

A structural VAR model describes a vector autoregressive process with parameters that offer behavioral interpretations. Although equation (1) is sufficient to define a VAR model, it describes variables as a function of only past values of the variables and error terms. A structural VAR model sometimes provides more substantively meaningful parameters because it describes variables as a function of contemporaneous values as well as past values of the variables.

A structural VAR model uses the more general parameterization given by

$$\mathbf{A}_0\mathbf{y}_t = \mathbf{a} + \mathbf{A}_1\mathbf{y}_{t-1} + \ldots + \mathbf{A}_p\mathbf{y}_{t-p} + \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0, \mathbf{U}) \tag{2}$$

with the diagonal elements of $\mathbf{A}_0$ typically assumed to be equal to 1 and $\mathbf{U}$ typically assumed to be diagonal. Off-diagonal elements of $\mathbf{A}_0$ capture contemporaneous relationships between the processes. The parameters of a structural VAR model are unidentified unless further restrictions are imposed. The structural model is equivalent to the model given by equation (1) with $\mathbf{B}_i = \mathbf{A}_0^{-1}\mathbf{A}_i$, $\boldsymbol{\alpha} = \mathbf{A}_0^{-1}\mathbf{a}$, and $\mathbf{T} = \mathbf{A}_0^{-1}\mathbf{U}\left(\mathbf{A}_0^{-1}\right)^{\top}$. This description is referred to as the reduced form of the model.

2.2 *Estimating Public Opinion*

Several methods have also been developed to construct time series of public opinion from polling data. The most frequently applied of these is the dyad ratios algorithm developed by

[3] When discussing a VAR(1) model, this article will refer to the matrix $\mathbf{B}_1$ as $\mathbf{B}$. This allows the notation for an element of this matrix to be simplified to $\boldsymbol{B}_{i,j}$.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

Stimson (1999). This algorithm and its software implementation WCALC (Stimson 2008)[4] seek to address irregular spacing and measurement error as well as allowing multiple series of polls into a measure of a single latent construct.[5] An alternative method for estimating public opinion is the Samplemiser method developed by Green et al. (1999) and based around a Kalman filter (see, e.g., Beck 1989). Samplemiser also allows for irregular spacing and measurement error, although it does not combine multiple series into a single measure.

2.3 *Temporal Aggregation Bias*

Regularly spaced time series data are sometimes formed by averaging or summing over intervals of time, known as temporal aggregation, or using only observations taken at fixed intervals of time, known as systematic sampling. This can occur for two different reasons. First, instantaneous measurement may be impossible or the only data available to researchers may already have been aggregated or sampled. Second, researchers may aggregate or sample the data themselves in order to make the data suitable for a model that does not allow irregularly spaced or mixed-frequency observations. Some data, such as survey marginals from a series of polls, are both irregularly spaced and already temporally aggregated.

Temporal aggregation can invalidate not only point estimates but also inferences about the direction or directions of causality (Freeman 1989), such as turning a unidirectional causal relationship into a bidirectional causal relationship (Tiao and Wei 1976). Temporal aggregation and systematic sampling also have other negative consequences. By ignoring information in the data, they lead to larger standard errors and lower statistical power. Different choices can also lead to different inferences when many possible choices exist for the aggregated time units (Shellman 2004). Several studies (e.g., Freeman 1989; Shellman 2004; Streb et al. 2012) have demonstrated that temporal aggregation can significantly impact the conclusions drawn from analyses of real political data.

2.4 *Imputation Bias*

Temporal aggregation and systematic sampling are not the only methods used when applying a model that assumes regular spacing to irregularly spaced data. With polling data in particular, another solution is common: imputation. Researchers interested in the relationship between public opinion and other variables often use WCALC or Samplemiser to construct a time series for public opinion that is then used as data in a VAR or other time series model.[6] Although this imputation process can reduce temporal aggregation bias by permitting shorter time units, it also introduces other biases.

The use of imputed series as data can add bias in two different ways. First, treating estimated values of unobserved public opinion as the observed values amounts to best-guess or conditional-mean imputation. This leads to invalid results because it ignores the uncertainty in the estimated values and the correlation between them. More importantly, the imputation model is not the correct model for the data. Since public opinion is imputed based only on polls, the imputation model implicitly treats public opinion as unrelated to any other series. If imputed data is then used to test the relationship between public opinion and another series, this implicit assumption in the imputation process biases our tests toward the conclusion that the two are unrelated.

2.5 *State-Space Models*

State-space models are a class of models that explicitly model observations as a function of latent processes. Various types of state-space models have gained wide usage in engineering, biology, physics, and economics. However, state-space models have garnered relatively little usage in

[4]The term WCALC also, sometimes, seems to be used as a synonym for the dyad ratios algorithm. For the sake of brevity, it will be used that way in this article.

[5]Users still need to select a time unit for aggregation—WCALC offers daily, monthly, quarterly, annual, or multiple years as options—but time intervals that include no polls are permissible.

[6]For recent examples, see, e.g., Carlin et al. (forthcoming) and Enns (2014).

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

political science.[7] A state-space model consists of a transition equation and an observation equation.[8] The transition equation captures the evolution of the state variable over time but is not directly observed. The observation equation links observations to the state variable.[9]

A state-space model in which the latent process is a vector autoregressive process and observations are linear functions of that process with noise is known as a multivariate autoregressive state-space (MARSS) model.[10] MARSS models extend VAR by allowing for measurement error as well as missing observations. A longer description of MARSS models and their relationship to the model developed in this article can be found in the Appendix available as Supplementary Material.

2.6 *Continuous-Time Models*

A continuous-time process is a time-varying process that is indexed by continuous, rather than discrete, values of time. Discrete-time models are sometimes applied to data from continuous-time processes, but this simplification comes at a cost of reduced accuracy.[11] A continuous-time model allows us to avoid this cost.

Within political science, continuous-time models have been almost entirely limited to event history analysis, in which durations are often treated as a continuous quantity (Box-Steffensmeier and Jones 1997, 2004), and longitudinal network analysis (e.g., Berardo and Scholz 2010; Manger et al. 2012), with a very small number of additional uses in formal models (e.g., Fearon 1994, 1998; Carpenter 2003). However, they have played an important role in physics for over a century and have also seen use in chemistry, biology, and engineering (Hänggi and Marchesoni 2005). Their use has also become common in finance—as demonstrated by the ubiquitous Black–Scholes option-pricing model—and applications have become increasingly common in econometrics (Brockwell 2001).

Continuous-time models have proved useful in analyzing data with irregular spacing. They are particularly advantageous when the data lack an underlying sampling interval (Jones 1985). Although much of this work has focused on the univariate case (Jones 1981, 1985; Jones and Tryon 1987; Jones and Ackerson 1990), some has also addressed the multivariate case (Jones 1984; Jones and Boadi-Boateng 1991).

Continuous-time models are not the only approach to analyzing irregularly spaced data. Another approach, most relevant when irregularly spacing occurs because of missingness in otherwise regularly spaced observations, has been to extend methods for regularly spaced time series to allow for missing observations. Other literatures have addressed mixed-frequency data and the frequency spectrum of irregularly spaced data. A brief overview of these other approaches as well as a longer discussion of continuous-time models appears in the Appendix available as Supplementary Material. However, beyond dealing with irregularly spaced data, most of these literatures are aimed at addressing different issues than those discussed in Section 1.1.

3 The MONOCAR Model

This section presents a new model that addresses the shortcomings of the VAR model while keeping the model as similar to a VAR model as possible. The resulting model is referred to as the multivariate overlapping, noisy observation continuous-time autoregressive (MONOCAR) model. Although this is not the first multivariate continuous-time autoregressive (MCAR) model, it is,

[7] Important exceptions include Beck (1989); Green et al. (1999); McAvoy (1998); Kellstedt et al. (1993); Brandt et al. (2000); Brandt and Williams (2001); Martin and Quinn (2002); Bond et al. (2003); Jackman (2005); Pickup and Johnston (2007); De Boef and Keele (2008); Armstrong (2009).

[8] See Commandeur and Koopman (2007) for an introduction to state-space models and Durbin and Koopman (2012) for a more advanced treatment.

[9] When the latent variables are discrete, the models are more commonly known as hidden Markov models.

[10] The terms vector autoregressive, state-space model, or dynamic linear model are used in various contexts.

[11] Continuous-time processes and discrete-time processes are heavily related. Discrete-time processes can often be represented as regularly spaced observations of continuous-time processes and some textbooks introduce time series models in a general framework that includes continuous-time models before focusing on discrete-time models (e.g., Fuller 1976). However, continuous-time models are most commonly applied when discrete-time representations are not possible.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

to the author's knowledge, the first to allow noninstantaneous, overlapping observations with measurement error of an MCAR or other similar process. Allowing observations to have collection times that overlap but do not entirely coincide seems to be the most novel feature of the model and captures an important characteristic of polls and many other types of data.

3.1 *Latent Process*

The model can be split into an unobserved latent process and an observation process that relates this latent process to the data. To avoid the need for regularly spaced observations, the latent process is assumed to be a continuous-time process. MONOCAR assumes the latent process is a multivariate Ornstein–Uhlenbeck (OU) process defined by the stochastic differential equation

$$d\mathbf{y}_t = \boldsymbol{\Theta}\big(\boldsymbol{\mu} - \mathbf{y}_t\big)dt + \Sigma^{\frac{1}{2}} d\mathbf{W}_t, \tag{3}$$

where $\boldsymbol{\Theta}$ is a matrix with no negative eigenvalues, Σ is a positive-semidefinite matrix, and $\mathbf{W}_t$ is a multivariate Wiener process.

A multivariate OU process is a continuous-time analogue of a first-order Gaussian vector autoregressive process. Like a VAR(1) process, an OU process is a Gaussian process and a Markov process. Both are also stationary processes for appropriate parameter values. The multivariate OU process also includes a Gaussian VAR(1) process as a special case. If the observations of a multivariate OU process occur only at integer values of time, the result is a VAR(1) process where $\boldsymbol{\alpha} = (\mathbf{I} - e^{-\boldsymbol{\Theta}})\boldsymbol{\mu}$, $\mathbf{B} = e^{-\boldsymbol{\Theta}}$, and $\mathrm{vec}\,(\mathbf{T}) = (\boldsymbol{\Theta} \oplus \boldsymbol{\Theta})^{-1}(\mathbf{I} - e^{-(\boldsymbol{\Theta}\oplus\boldsymbol{\Theta})})\,\mathrm{vec}\,(\boldsymbol{\Sigma})$.[12]

3.2 *Observation Process*

MONOCAR assumes that this multivariate OU process is a latent process that need not be fully observed. Rather, observations are noisy measurements of this latent process. For observations recorded over a nonzero length of time, the expected value of the observation is the mean of the latent process over that length of time. Thus, for a given observation, $z_{i,j}$, of variable i measured from time $t_{i,j,0}$ to $t_{i,j,1}$,

$$z_{i,j} = \frac{1}{t_{i,j,1} - t_{i,j,0}} \int_{t_{i,j,0}}^{t_{i,j,1}} y_{i,s} ds + \eta_{i,j}, \quad \eta_{i,j} \sim \mathcal{N}\Big(0, \psi_{i,j}^2\Big) \tag{4}$$

with $\psi_{i,j}^2$ representing the known or parametrized variance of the observation $z_{i,j}$ conditional on the values of $\mathbf{y}_t$. This can easily be extended to allow multivariate observations whose expectations are linear transformations of the average of $\mathbf{y}_t$. In the case where observation is instantaneous ($t_{i,j,0} = t_{i,j,1}$),

$$z_{i,j} = y_{i,t_{i,j,0}} + \eta_{i,j}, \quad \eta_{i,j} \sim \mathcal{N}\Big(0, \psi_{i,j}^2\Big). \tag{5}$$

Perfect measurements of the latent process are possible by setting $\psi_{i,j}^2 = 0$.

The MONOCAR model, therefore, allows for observations that are noisy and neither instantaneous nor regularly spaced. It can also include a VAR(1) model as a special case where $t_{i,j,0} = t_{i,j,1} = j$ and $\psi_{i,j}^2 = 0$ for all i and j, meaning simply that the data consist of instantaneous observations measured every one unit of time without noise.

[12]The symbol $\otimes$ indicates the Kronecker product, defined as $\mathbf{X} \otimes \mathbf{Y} = \begin{pmatrix} X_{11}\mathbf{Y} & \dots & X_{1n}\mathbf{Y} \\ \vdots & \ddots & \vdots \\ X_{m1}\mathbf{Y} & \dots & X_{mn}\mathbf{Y} \end{pmatrix}$; the symbol $\oplus$ indicates the Kronecker sum, defined as $\mathbf{X} \oplus \mathbf{Y} = \mathbf{X} \otimes \mathbf{I}_{pq} + \mathbf{I}_{mn} \otimes \mathbf{Y}$; and vec indicates matrix vectorization, defined as $\mathrm{vec}\,(\mathbf{X}) = (X_{11}, \dots, X_{m1}, \dots, X_{1n}, \dots, X_{mn})^{\mathrm{T}}$, where $\mathbf{X}$ is the $m \times n$ matrix $\begin{pmatrix} X_{11} & \dots & X_{1n} \\ \vdots & \ddots & \vdots \\ X_{m1} & \dots & X_{mn} \end{pmatrix}$, $\mathbf{Y}$ is a $p \times q$ matrix, and $\mathbf{I}_{mn}$ is the $m \times n$ identity matrix.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

3.3 *Variance-Stabilizing Transformations*

An additional complication occurs when the variance of each observation is dependent on its mean. Observations following a binomial distribution with mean π determined by the latent process, as might be assumed with polling data, would lead to a variance of $\frac{\pi(1-\pi)}{N}$. Variance-stabilizing transformations provide a method of dealing with this complication by transforming the data to produce observations that can be well approximated by a Gaussian distribution with constant variance.

The primary motivation for doing so is to ensure that the variance is (approximately) known or, at least, not dependent on the latent process. These transformations also typically have the side benefit of making the resulting distribution even closer to a Gaussian. When the variance of each observation does not depend on the latent process itself, no variance-stabilizing transformation is needed.

There are also many alternatives to the use of variance-stabilizing transformations when the variance does depend on the latent process that might be used in their place, although the alternatives have their own complications.

Two common cases that would require variance-stabilizing transformations are the binomial and Poisson distributions. The transformations

$$z' = \arcsin\left(\left(1 - \frac{1}{5000}\right)\left(2\frac{z}{N} - 1\right)\right) \tag{6}$$

for the binomial distribution and

$$z' = \sqrt{z + \frac{3}{8}} \tag{7}$$

for the Poisson distribution are possible variance-stabilizing transformations for these cases. Their respective approximate variances are $\frac{1}{N}$ and $\frac{1}{4}$.

3.4 *Additional Variance Concerns*

Independent of the need to stabilize the variance, Pickup and Wlezien (2009) argue that the variance implied by the sample size of a poll may differ from the true variance as a result of weighting and other common polling practices. Pickup and Johnston (2008) recommend adding a variance multiplier and an additional component to the variance of each observation of the form

$$\mathrm{Var}\left[z_{i,j}\right] = \gamma_i + \nu_i \psi_{i,j}^2, \tag{8}$$

where $\psi_{i,j}^2$ is the untransformed variance implied by the sample size, ν_i is a parameter multiplying the untransformed variance and γ_i is a parameter adding additional variance.

This, however, can suggest negative variances for certain parameter values.[13] An appealing alternative is to assume that

$$\mathrm{Var}\left[z_{i,j}\right] = (\nu_i)\left(\psi_{i,j}^2\right)^{\xi_i}, \tag{9}$$

where ν_i is a parameter multiplying the untransformed variance and ξ_i is a parameter exponentiating the untransformed variance.

This approach eliminates the possibility of negative variances so long as ν is constrained to be positive.

In some data sets, the variance is constant or nearly constant within each series. An example of this can be seen in polling data when all or nearly all polls have the same sample size and a variance-stabilizing transformation is used. In these cases, it may be simpler to use only a single

[13] Restricting the parameter values to ensure this does not happen complicates inference in a maximum-likelihood framework when the estimates are at these bounds. This is not problematic in a Bayesian framework, as is applied by Pickup and Johnston (2008) to models with this variance transformation.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

parameter to adjust the variances—for example, to use a variance multiplier, ν, with γ and ξ above implicitly set to zero and one, respectively.

3.5 *Other Covariates*

In discrete-time models, other covariates are often included as exogenous variables. This is more complicated in a continuous-time setting. Because the latent process in the MONOCAR model takes on values at infinitely many points in time, it is not possible to store values for covariates at all these points in time. Nonetheless, MONOCAR can include other covariates through several approaches for different types of variables.

3.5.1 Observation equation

Variables that are tied to observations are straightforward additions to the MONOCAR model by including an additional term that is a linear (or even nonlinear) function of the covariates to the expectations of the observations. Thus, the observation process could become

$$z_{i,j} = \mathbf{x}_j{}^{\top}\boldsymbol{\zeta}_i + \frac{1}{t_{i,j,1} - t_{i,j,0}} \int_{t_{i,j,0}}^{t_{i,j,1}} y_{i,s} ds + \eta_{i,j,} \tag{10}$$

where $\mathbf{x}_j$ is a vector of covariates for observation j and $\boldsymbol{\zeta}_i$ is a vector of coefficients for series i.

One example of a variable that is clearly tied to the observation is a dummy variable for the organization that conducted a poll. Different polling organizations typically use slightly different methodologies, leading to systematic biases in their results. The inclusion of a term in the observation equation for the bias of each polling organization, termed a "house effect," is a common solution to this problem (see Jackman 2005).[14]

3.5.2 Functions of time

Some variables can be known at all points in time without needing to store values at all points in time because they can be expressed as a function of time. In these cases, a time-varying function that depends on these variables can replace the constant parameter μ, representing the long-run mean of the latent process. If $x(t)$ is such a variable, μ can be replaced by $\mu(t) = \boldsymbol{\kappa}_0 + \boldsymbol{\kappa}_1 x(t)$, where $\boldsymbol{\kappa}_0$ and $\boldsymbol{\kappa}_1$ are new parameters. The latent process given by equation (4) can then be replaced by

$$d\mathbf{y}_t = \boldsymbol{\Theta}\big(\boldsymbol{\kappa}_0 + \boldsymbol{\kappa}_1 x(t) - \mathbf{y}_t\big)dt + \Sigma^{\frac{1}{2}} d\mathbf{W}_t. \tag{11}$$

For most functions $x(t)$, computing the likelihood under such a model might be difficult. However, if $x(t)$ is a piecewise-constant function of time, the computation becomes much simpler. In this case, although the latent process no longer follows an OU process in total, it does follow an OU process over each interval of time for which $x(t)$ is constant. As a result, only minor changes to the likelihood computation are needed.

3.5.3 Latent variables

Variables can also be included as additional components to the latent process. The variable is treated as being only partially observed. The variable may be known precisely at some points in time, but is not precisely known at other points in time. Observations can also include measurement error or be recorded over intervals of time, in which case the latent variable is never known precisely.

[14]More discussion of house effects appears in the Appendix available as Supplementary Material.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

If a latent series is exogenous, the parameters representing the effect of other latent series on it and the correlation of the error process with the error processes of other series are zero. If an exogenous variable corresponds to component j of the latent process, then $\mathbf{\Theta}_{j,k} = 0$, $\Sigma_{j,k} = 0$, and $\Sigma_{k,j} = 0$ for all $k \neq j$. These parameter restrictions allow for more efficient estimation of the model.

3.6 *Structural Form*

A more general parameterization of the MONOCAR model may provide a behavioral interpretation of the parameters, analogous to the structural form of a VAR. The structural form of a MONOCAR differs from the reduced form by capturing nongradual contemporaneous effects. Unlike the reduced form of a VAR model, the reduced form of a MONOCAR model already treats each series as potentially having a contemporaneous effect on the other series. However, these effects are gradual; while there is no lag before a shock to one series begins to affect another series, the size of this effect grows with time.

A structural MONOCAR model adds contemporaneous effects in which the full effect of a shock to one series appears immediately in another series. This model is constructed by replacing equation (4) with

$$\mathbf{H}d\mathbf{y}_t = \mathbf{\Theta}\big(\boldsymbol{\mu} - \mathbf{y}_t\big)dt + \Sigma^{\frac{1}{2}}d\mathbf{W}_t. \tag{12}$$

The reduced form of this model is

$$d\mathbf{y}_t = \mathbf{H}^{-1}\mathbf{\Theta}\big(\boldsymbol{\mu} - \mathbf{y}_t\big)dt + \mathbf{H}^{-1}\Sigma^{\frac{1}{2}}d\mathbf{W}_t. \tag{13}$$

As with structural VAR, the structural MONOCAR model is unidentified without restrictions on the parameters in equation (13).

3.7 *Estimation*

Estimation and inference for the parameters of the MONOCAR model can be performed in a maximum-likelihood framework. An exact likelihood of a MONOCAR model can be computed by building upon a multivariate Kalman filter (Schweppe 1965).[15] Further details of the likelihood computation can be found in the Appendix available as Supplementary Material. An R package implementation is also available.

An alternative approach is to perform estimation and inference in a Bayesian framework with Markov chain Monte Carlo (MCMC) methods.[16] A Bayesian approach to MONOCAR has several advantages, such as avoiding the need to rely on asymptotic results for inference and allowing prior information to be included.[17] On the other hand, the maximum-likelihood approach offers estimates to be computed much more quickly, allows inference to be performed in a frequentist framework, and avoids the need to specify prior distributions for parameters.

Unit roots arise in a MONOCAR model when an eigenvalue of $\mathbf{\Theta}$ has a real part 0. As with the least-squares estimator of a VAR process, unit roots can bias the estimates of a MONOCAR model. Throughout this article, the underlying OU process is assumed to be stationary, and therefore the unit-root problem to be inapplicable. Extensions of the MONOCAR estimator to ensure consistency in the presence of a unit root or formal hypothesis tests of unit roots form interesting topics for future research but are beyond the scope of this article.

[15] For an overview of the Kalman filter, see Beck (1989).
[16] A more detailed discussion of Bayesian estimation of MONOCAR appears in the Appendix available as Supplementary Material.
[17] Many advantages offered by a Bayesian approach to MONOCAR are similar to those discussed with other time series models (see, e.g., Brandt and Freeman 2006, 2009). Priors may also be particularly useful in dealing with the variance parameters described in Section 3.4.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

3.8 *Interpretation*

The coefficients of the MONOCAR model can be interpreted in a manner relatively similar to a VAR model. First, $\boldsymbol{\mu}$ is perhaps easiest to interpret, as it represents the mean of the stationary distribution (or long-run mean) of the latent process, $\mathbf{y}$. The variance–covariance matrix $\boldsymbol{\Sigma}$ is also fairly easy to interpret, as the diagonal elements provide the variance of the exogenous movement in each series and the off-diagonal elements provide the covariance between the series that is unexplained by the modeled effect each series has on the other.

The matrix of coefficients, $\boldsymbol{\Theta}$, is slightly more difficult to interpret. The diagonal elements represent the rate at which each series reverts to its mean (absent the effects of the other series). A value of zero would indicate that a series has no mean-reversion tendency on its own (similar to a coefficient of one in a VAR model).[18] The higher the value of a diagonal element, the faster the effects of a shock to the series disappear.

Off-diagonal elements of $\boldsymbol{\Theta}$ are similar to their VAR counterparts, albeit with a sign flip. A positive value for $\Theta_{i,j}$ indicates that when series j is large (relative to μ_j), it tends to make series i smaller. Thus, for example, if a high level of one series tends to lead a second series to increase (until it reaches a new, higher equilibrium level), one would expect a *negative* coefficient on the relevant element of $\boldsymbol{\Theta}$.

4 Monte Carlo Studies

This section illustrates some of the issues discussed in this article through two Monte Carlo studies. The first study addresses temporal aggregation. The second addresses irregularly spaced observations and the consequences of imputation. In each case, the simulated data are analyzed with a standard technique and with MONOCAR.

These studies focus on the relationship between two series. In all cases, the second series is assumed to affect the first series but not the reverse. The analyses will focus on the estimated effect of the first series on the second.[19] To simplify the presentation, only a single set of parameters of the latent process is included.[20] Additional Monte Carlo studies that cover a wider range of parameters and comparisons to other models appear in the Appendix available as Supplementary Material.

Each study will focus on estimates of the effect of the first series on the second series and a test of the hypothesis that this effect is zero. Since the data-generating process will assume that no such effect exists, an unbiased estimator should produce estimates with mean zero. Further, the p-value for an unbiased hypothesis test should follow (or at least be stochastically dominated by) a uniform distribution. As a result, a quantile–quantile plot of the simulated distribution of p-values against a uniform distribution should lie on a 45° line. If the points fall below a 45° line, the null hypothesis is being erroneously rejected at a greater rate than the nominal significance level of the test.

4.1 *Temporal Aggregation*

4.1.1 Overlapping observations

Temporal aggregation is known to cause bias in VAR models and overlapping observation intervals have the potential to increase the bias even more. This section presents a Monte Carlo study in

[18] A coefficient of zero does not necessarily imply that a series is not mean-reverting once effects of the other series are taken into account.

[19] This simplifies interpretation because, under either parameterization, the relevant parameter (either $\mathbf{B}_{2,1}$ or $\boldsymbol{\Theta}_{2,1}$) is zero if, as is true for the data-generating process, the first series does not affect the second. When looking at the distribution of the estimates from the MONOCAR model, the parameters are transformed to use the VAR parameterization for ease of comparison.

[20] Specifically, the simulations will assume that the latent process follows a multivariate OU process with $\boldsymbol{\Theta} = \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{pmatrix}$ and $\boldsymbol{\Sigma} = \mathbf{I}$.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

which observations are temporally aggregated into overlapping or nonoverlapping time intervals. This study demonstrates bias in VAR estimates and compares with estimates of MONOCAR models.

Observations are assumed to begin at integer values of time. The length of the observation intervals varies from one, implying no overlap between observations, to three, implying that two-thirds of the time interval for one observation overlaps with the time interval for the next. In each simulation and for each estimator, the estimated effect of the first series on the second series is computed along with a p-value for a test of the hypothesis that this effect is zero.

Figure 1 shows the distribution of estimates and quantile–quantile plot of p-values against a standard uniform. For the VAR models, we see a noticeable negative bias in the estimates seen in the previous study even when the intervals are consecutive. As the amount of overlap between observation intervals increases, the size of this bias increases. A similar pattern is observed in the p-values. Although the estimates and p-values are biased even when no overlap is present, a 50% overlap between observations creates a situation in which virtually all estimates are negative and the null hypothesis is erroneously rejected in most cases.

MONOCAR models exhibit no bias regardless of the amount of overlap between observations. Similarly, the quantile–quantile plots show that the p-values from the MONOCAR models follow the desired uniform distribution that should be expected under the null hypothesis. Thus, while temporal aggregation poses an increasing problem for VAR models as the amount of overlap between time intervals increases, MONOCAR continues to produce reliable estimates and inference.

4.2 *Imputation*

The second Monte Carlo study deals with bias from imputation. This study will test this approach on simulated, irregularly spaced data and compare the resulting estimates and p-values with those from MONOCAR. Imputation bias is not an issue with approaches such as MARSS models, which can deal effectively with irregularly spaced data in discrete-time, although temporal aggregation and overlapping observations can still introduce bias. However, because the WCALC-imputation approach is often taken with polling data, it is worth analyzing the bias that results from this approach.

The simulated latent processes are identical to those used in the first Monte Carlo study, but the observation process differs. Observations are instantaneous measures of the latent process over a forty-year period. Observations of the first series, which can be thought of as simulated polls, are collected at random points over this period. Observations of the second series are regularly spaced and occur at the midpoint of each month. Measurement error is also added to observations from both series.[21] Thus, temporal aggregation is not a problem with the simulated data, but irregular spacing and noise are.

Estimates and hypothesis tests are then computed with two different methods. In the first method, WCALC is applied to the observations of the first series.[22] Estimates are formed from a VAR model applied to the imputed monthly observations for the first series paired with the corresponding monthly observations of the second series. These results are compared with estimates of MONOCAR models based on the original simulated data without imputation. Histograms of estimates and quantile–quantile plots of p-values of both models are shown in Fig. 2.

The estimates and p-values from WCALC and VAR show clear evidence of bias. The mean of the estimates is noticeably greater than the true value of zero, with few estimates lying at or below zero. Similar issues arise with the p-values. The p-values of an unbiased hypothesis test should lie on

[21]The measurement error consists of independent draws from a normal distribution with a mean of zero and standard deviation of 0.01 that are added to each observation.

[22]Monthly time units are used to match the frequency of the second series. Using Samplemiser over WCALC leads to quite similar results in the Appendix available as Supplementary Material.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

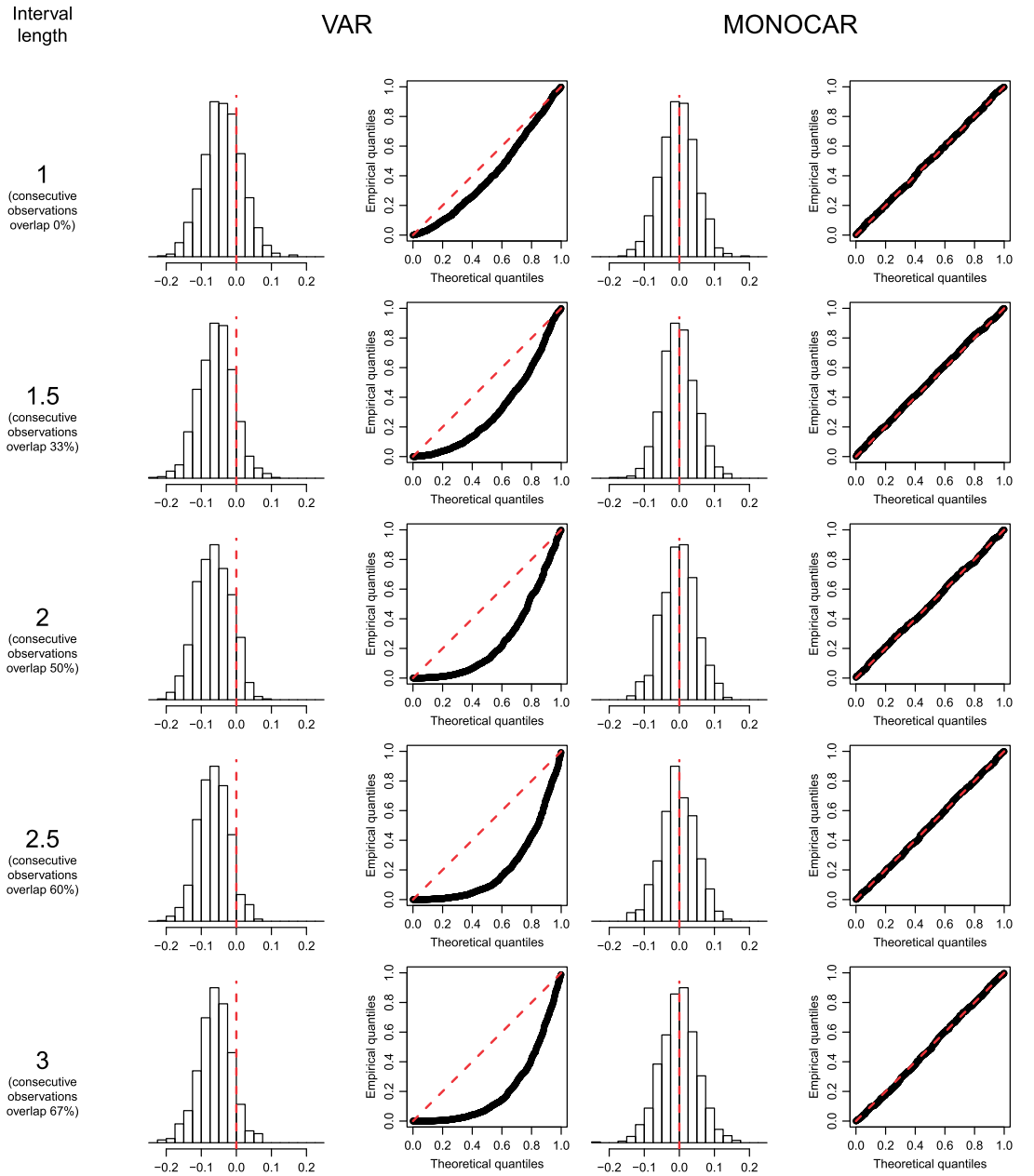


Fig. 1 Histogram of distribution of estimates and quantile–quantile plot of *p*-values from VAR and MONOCAR models for the overlapping observation Monte Carlo study. 1,000 simulations per study. In each simulation, the data consist of 1,000 observations per series. MONOCAR estimates are transformed to use the same parameterization as the VAR model. The estimated parameter shown is $\mathbf{B}_{1,2}$, which equals zero under the data-generating process. A dashed line shows this true value of $\mathbf{B}_{1,2} = 0$ in each histogram and a 45° line in each quantile-quantile plot. The *p*-values correspond to a test of the null hypothesis that $\mathbf{B}_{1,2} = 0$. Since the null hypothesis is true under the data-generating process, valid *p*-values should follow a uniform distribution.

a 45° line, but the *p*-values from WCALC and VAR lie well below the line, indicating that the null hypothesis is erroneously rejected at a far greater rate than the nominal level of the test.

In contrast, MONOCAR again shows no signs of bias in either estimation or inference. As can be seen in the right half of Fig. 2, the distribution of the estimates appears centered around the true

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

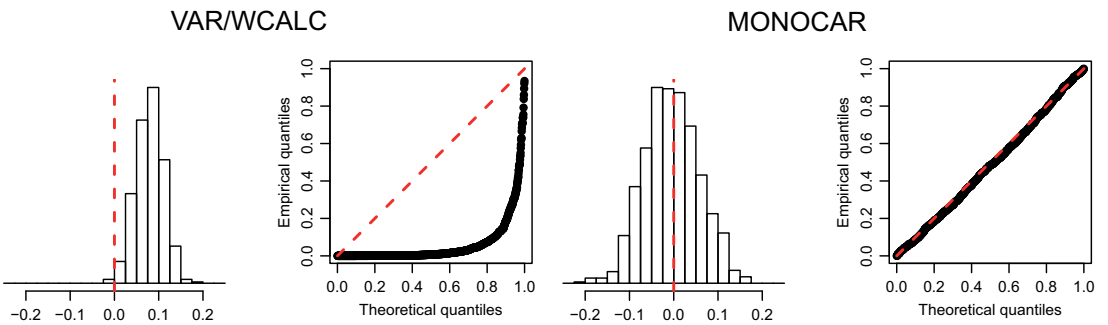


Fig. 2 Histogram of distribution of estimates and quantile–quantile plot of *p*-values from VAR/WCALC and MONOCAR models for the imputation Monte Carlo study. 1,000 simulations per study. In each simulation, the first series consists of 320 randomly spaced instantaneous polls (averaging eight per year) and the second series consists of monthly instantaneous observations. The estimated parameter shown is $\mathbf{B}_{1,2}$, which equals zero under the data-generating process. A dashed line shows this true value of $\mathbf{B}_{1,2} = 0$ in each histogram and a 45° line in each quantile-quantile plot. The *p*-values in the quantile–quantile plots correspond to a test of the null hypothesis that $\mathbf{B}_{1,2} = 0$. Since the null hypothesis is true under the data-generating process, valid *p*-values should follow a uniform distribution.

value of zero. Moreover, the quantile–quantile plot shows that the *p*-values follow a uniform distribution. Since the null hypothesis is true by assumption, this is precisely what we should observe from an unbiased test.

Thus, the common practice of using estimates of public opinion from WCALC or Samplemiser as data in a time series analysis can be unreliable. Even if this process avoids temporal-aggregation bias, it does so at the cost of creating a new source of bias. However, by avoiding the need to aggregate or impute data, MONOCAR can produce unbiased estimates and inference.

The bias illustrated in this study does not mean that WCALC, Samplemiser, or other such techniques are unreliable for other purposes. As estimators of the latent series themselves, they often perform quite well—including on the simulated data in this study. Rather, it is the use of estimated latent series as data that leads to the biases seen here.

5 Applications

This section provides two examples of practical applications of MONOCAR. In addition to demonstrating the results and interpretation of the MONOCAR model, these applications also allow us to compare these results of a more typical use of a VAR model. The goal of this section is not to provide a full exploration of the relationships in question, which would require a much more complete analysis, but rather to demonstrate the usefulness of the method to practical applications and real data.

These two examples demonstrate applications of MONOCAR to two quite different types of data and subfields of political science. The first example shows an application to event data common in international relations. The second shows an application to public opinion data and compares results from MONOCAR with the common combination of WCALC and VAR models.

5.1 *Bosnia*

Goldstein and Pevehouse (1997) provide an excellent example of the use of events data with VAR models. In the article, the authors analyze the role of reciprocity in cooperation during the conflict in Bosnia from 1992 to 1995. Using data on the level of net cooperation and VAR models, Goldstein and Pevehouse test for two different types of responses: bilateral responses, in which the net cooperation of one actor toward another affects the net cooperation of the second actor toward the first; and triangular responses, in which the net cooperation of one actor toward another

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

affects the net cooperation of the second actor toward a third actor or the net cooperation of a third actor toward the first.

Goldstein and Pevehouse (1997) face the problem of irregularly spaced observations in their data. Although daily observations were collected, many days contain no observations. The authors address these issues by aggregating the daily observations into weeks and by coding the small percentage of weeks without a news report as a zero.

5.1.1 Data

This analysis uses the same data collected by Goldstein and Pevehouse using the Kansas Event Data System, the World Event Interaction Survey (WEIS) coding scheme, and the Goldstein (1992) net-cooperation scale. Each observation provides an interval measure of the level of net cooperation for a given dyad at a given point in time, with positive values indicating cooperation and negative values indicating conflict. As in the original study, the data are divided into three periods of time: a "drift" period from April 17, 1993, to February 12, 1994; a "threats" period from February 12, 1994, to December 17, 1994, and a "promises" period from December 17, 1994, to July 29, 1995.

The models again parallel those used by Goldstein and Pevehouse. Separate models are run for each period. Models of the first two periods include three series: the international community's actions toward Serbia (*IS*), Serbian actions toward the international community (*SI*), and Serbian actions toward Bosnia (*SB*). The model for the third period adds a fourth series capturing net cooperation of Bosnian actions toward Serbia (*BS*).

The MONOCAR analysis follows the original Goldstein and Pevehouse analysis with several important differences. Rather than aggregating into weekly totals, the MONOCAR analysis uses daily observations. Periods with no news reports in the data are treated as unobserved rather than as having a level of net cooperation of zero. Observations are treated as noisy measures of the level of net cooperation for the relevant dyad, averaged over the day. As a result of temporal aggregation in the VAR models, the sample sizes for the VAR models are smaller than for the MONOCAR models. This reflects the ability of MONOCAR to use the more fine-grained data despite irregular spacing.

5.1.2 Results

The VAR models provide an exact replication of the results found in Goldstein and Pevehouse (1997).[23] These are presented in the left half of Table 1. Three coefficients of interest are statistically significant, but each only in a single period. First, the effect of Serbian actions toward Bosnia (*SB*) on international actions toward Serbia (*IS*) is significant in the second period. Second, the effect of international actions toward Serbia (*IS*) on Serbian actions toward the international community (*SI*) is significant in the third period. Finally, the effects of Serbian actions toward Bosnia on Bosnian actions toward Serbs is also statistically significant in the third period. In all cases, these same coefficients are not significant in models for either of the other two periods.

The MONOCAR estimates presented in the right half of Table 1 show somewhat different results. Notably, the same coefficients are significant with the same signs in periods two and three, whereas each was only significant in a single period before. The effect of Serbian actions toward Bosnia (*SB*) on international actions toward Serbia (*IS*) is significant and positive in both the second and third period, strengthening the case for a triangular response by the international community to Serbian actions toward Bosnia. Similarly, the effect of international actions toward Serbia (*IS*) on Serbian actions toward the international community (*SI*) is also significant in the second and third periods, strengthening the case for bilateral reciprocity in Serbian behavior toward the international community.

[23]Specifically, these estimates duplicate those in Table 3 of Goldstein and Pevehouse (1997).

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

Table 1 Models of Bosnia data

	VAR			*MONOCAR*		
	Period 1 (Drift)	*Period 2 (Threats)*	*Period 3 (Promises)*	*Period 1 (Drift)*	*Period 2 (Threats)*	*Period 3 (Promises)*
Explaining International toward Serbia						
IS ←*IS*	0.14	−0.01	0.24	59.09*	7.34*	7.56*
	(0.16)	(0.20)	(0.27)	(16.34)	(2.82)	(2.01)
IS ←*SI*	0.13	−0.59	−0.05	6.26	4.90	− 2.45
	(0.26)	(0.47)	(0.33)	(34.54)	(7.66)	(1.72)
IS ←*SB*	0.02	0.77*	0.12	−0.25	4.22*	3.99*
	(0.11)	(0.26)	(0.28)	(10.88)	(2.14)	(1.84)
IS ←*BS*			0.15			2.88
			(0.61)			(3.49)
IS ← constant/mean	−13.73*	−35.28	−22.02	−4.10*	−9.52*	−6.08*
	(6.24)	(21.72)	(20.10)	(0.72)	(3.11)	(1.90)
Explaining Serbia toward International						
SI ←*IS*	0.01	0.01	0.50*	3.45	0.61*	7.66*
	(0.09)	(0.09)	(0.16)	(6.89)	(0.27)	(1.82)
SI ←*SI*	−0.24	−0.21	0.07	25.54	3.13*	13.09*
	(0.15)	(0.21)	(0.20)	(22.83)	(1.55)	(3.24)
SI ←*SB*	0.09	0.24*	−0.17	2.78	0.29	−2.65
	(0.07)	(0.11)	(0.17)	(2.34)	(0.49)	(2.42)
SI ←*BS*			0.10			−2.63
			(0.36)			(5.79)
SI ← constant/mean	−9.31*	−21.13*	−29.53*	−2.98*	−5.63*	−7.60*
	(3.58)	(9.43)	(12.06)	(0.58)	(1.45)	(1.50)
Explaining Serbia toward Bosnia						
SB ←*IS*	−0.36	−0.23	0.19	−0.97	−0.37	−1.91
	(0.22)	(0.12)	(0.19)	(4.59)	(0.36)	(2.34)
SB ←*SI*	0.13	0.43	0.16	1.11	1.91	5.29
	(0.36)	(0.27)	(0.24)	(5.15)	(1.64)	(3.30)
SB ←*SB*	0.21	0.55*	0.24	13.46*	1.49*	14.45*
	(0.15)	(0.15)	(0.20)	(5.15)	(0.74)	(3.81)
SB ←*BS*			0.53			−1.32
			(0.44)			(5.02)
SB ← constant/mean	−27.37*	−27.25*	−19.23	−6.03*	−9.63*	−10.28*
	(8.38)	(12.52)	(14.51)	(1.07)	(2.60)	(1.45)
Explaining Bosnia toward Serbia						
BS ←*IS*			−0.03			−0.86
			(0.07)			(0.89)
BS ←*SI*			−0.04			0.69
			(0.09)			(1.43)
BS ←*SB*			0.25*			1.99
			(0.08)			(1.33)
BS ←*BS*			−0.25			9.27*
			(0.16)			(4.72)
BS ← constant/mean			−8.56			−3.64*
			(5.48)			(0.94)
N (IS)	43	44	32	185	253	170
N (SI)	43	44	32	149	205	154
N (SB)	43	44	32	197	227	148
N (BS)			32			108
log likelihood	−585.12	−729.88	−632.49	−1918.38	−2886.25	−2285.04

Note. Off-diagonal elements of Θ have been reversed in sign to ease interpretation. Coefficients are not directly comparable in magnitude between VAR and MONOCAR results but off-diagonal parameters are comparable in direction.
*$p < .05$.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

Overall, the MONOCAR results strengthen the original results of Goldstein and Pevehouse. The increased evidence for a triangular response by the international community may be of particular interest here. MONOCAR also shows a consistent pattern across periods two and three. And whereas the failure to find any coefficient of interest significant in more than one period might be viewed as cautionary, the consistency of finding identical patterns in periods two and three allows for greater confidence that the observed relationships are real.

5.2 *Presidential Approval, January 20, 1981, to January 20, 1993*

The second example looks at the relationship between the unemployment rate and presidential approval. The analysis focuses on the intuitive notion that higher unemployment leads to lower presidential approval using data from the presidencies of Ronald Reagan and George H. W. Bush. Several, though not all, studies have found evidence for this relationship.[24] The data are first analyzed by applying WCALC to the polling series and then estimating a VAR model using the resulting monthly measure of approval as data along with the unemployment rate.[25] The data are also analyzed with MONOCAR without using the imputed series from WCALC. In the MONOCAR analysis, polls are explicitly treated as noninstantaneous observations and the variance of each poll is also allowed to be a transformation of the sample-size-implied variance to address the critique of Pickup and Wlezien (2009), as discussed in Section 3.4.

5.2.1 Data

The data consist of the responses to presidential approval questions in polls conducted by nine organizations[26] during the Reagan and Bush administrations. The unemployment rate is the seasonally adjusted percentage of the labor force that is unemployed as measured by the Bureau of Labor Statistics.[27] Both VAR and MONOCAR models are estimated with and without dummy variables for the Bush presidency.[28]

5.2.2 Results

The results of applying WCALC and VAR to the data generally support a null finding.[29] The left half of Table 2 shows two models estimated using this approach, one with a dummy variable for the Bush administration included and one without. In both cases, we see no statistically significant effect of either series on the other.[30]

Applying MONOCAR provides a different inference. The right half of Table 2 shows these models, again estimated with and without a dummy variable for the Bush administration. The

[24]For a detailed discussion of existing research on the economic variables and presidential approval, see Berlemann and Enkelmann (2012).

[25]The use of WCALC requires the choice of time unit. Because the unemployment rate is only available on a monthly basis, the VAR/WCALC analysis aggregates the data into monthly observations.

[26]The questions are of the form "Do you approve or disapprove of the way . . . is handling his job as president?" The data include polls conducted by ABC News/Washington Post, CBS News/New York Times, the Gallup Organization, Gordon S. Black Corporation, NBC News/Wall Street Journal, Princeton Survey Research Associates, the Roper Organization, the Wirthlin Group, and Yankelovich Partners collected from the Roper Center's iPOLL Databank (Roper Center 2013).

[27]The data are from the Federal Reserve Economic Data (FRED) database maintained by the Federal Reserve Bank of St. Louis.

[28]This analysis does not include data on other factors that might affect presidential approval, although MONOCAR models implicitly capture factors that lead both series to move simultaneously by allowing the latent processes to be correlated.

[29]The use of Samplemiser in place of WCALC leads to similar results but would require assuming that there are no systematic differences between polls conducted by different organizations.

[30]All VAR models use a single lag. The use of BIC to select the number of lags leads to a single-lag model in all cases. The use of AIC to select the number of lags would lead to higher-order models. In all cases, a test of Granger causality is insignificant in the models selected using AIC.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

Table 2 Models of presidential approval and unemployment

	VAR/WCALC				*MONOCAR*			
	All polls		*Gallup only*		*All polls*		*Gallup only*	
Explaining presidential approval								
$A \leftarrow A$	0.892*	0.892*	0.899*	0.899*	0.302*	0.235*	0.245*	0.199*
	(0.036)	(0.036)	(0.035)	(0.036)	(0.082)	(0.073)	(0.081)	(0.075)
$A \leftarrow U$	−0.507	−0.494	−0.046	−0.045	−0.024*	−0.026*	−0.022*	−0.024*
	(0.292)	(0.310)	(0.025)	(0.026)	(0.011)	(0.011)	(0.010)	(0.010)
$A \leftarrow B$		0.107		0.007		−0.176*		−0.134
		(0.841)		(0.070)		(0.065)		(0.103)
$A \leftarrow$ *constant/mean*	9.692*	9.584*	0.322	0.314	0.126	0.186	0.117	0.162
	(3.623)	(3.734)	(0.179)	(0.198)	(0.076)	(0.097)	(0.079)	(0.097)
Explaining unemployment rate								
$U \leftarrow A$	0.001	0.001	0.020	0.017	0.041	0.072	0.136	0.151
	(0.001)	(0.001)	(0.017)	(0.017)	(0.102)	(0.085)	(0.107)	(0.091)
$U \leftarrow U$	0.997*	1.004*	1.000*	1.006*	0.011	0.007	0.003	−0.001
	(0.012)	(0.013)	(0.012)	(0.013)	(0.013)	(0.014)	(0.013)	(0.014)
$U \leftarrow B$		0.059		0.057		−0.219		−0.225
		(0.034)		(0.034)		(0.168)		(0.166)
$U \leftarrow$ *constant/mean*	−0.049	−0.109	0.001	−0.067	7.218*	7.301*	7.225*	7.326*
	(0.148)	(0.151)	(0.088)	(0.096)	(0.825)	(0.800)	(0.817)	(0.783)
N (approval)	144	144	144	144	533	533	212	212
N (unemployment)	144	144	144	144	144	144	144	144

Note: Number of approval observations reflects number of imputed observations from WCALC, not number of polls used by WCALC. Off-diagonal elements of $\boldsymbol{\Theta}$ have been reversed in sign to ease interpretation. Coefficients are not directly comparable in magnitude between VAR and MONOCAR but off-diagonal estimates are comparable in direction. See the Appendix available as Supplementary Material for full model estimates. A = presidential approval, U = unemployment rate, B = Bush administration.
*$p < .05$.

effect of unemployment on approval is statistically significant in both MONOCAR models. The effect of approval on unemployment is again insignificant.

Next, suppose only the data from a single polling organization had been collected. Table 2 also shows the results using only data from the most frequent polling organization in the data set, the Gallup Organization. Unsurprisingly, we see the same results: the effect of unemployment on presidential approval is insignificant in both models with VAR and WCALC. MONOCAR again leads to a different inference, with the effect of unemployment on presidential approval statistically significant in both models. Thus, WCALC and VAR do not show a statistically significant relationship with data from all nine polling organizations, whereas MONOCAR finds a statistically significant effect of unemployment on approval even when limited to only Gallup polls.

We reach a substantively different conclusion using MONOCAR than VAR and WCALC. In particular, results based on MONOCAR find evidence for a relationship in line with theoretical expectations, whereas results using VAR and WCALC do not. Moreover, the MONOCAR estimates show evidence of this relationship even when using substantially more limited data. Results based on MARSS models, which appear in the Appendix available as Supplementary Material, are more similar to MONOCAR than VAR models but still include important differences.[31] Although a more complete substantive analysis of the relationship would require further data and analysis,

[31]These MARSS models are as similar to the MONOCAR models in table 4 as possible, but treat time as occuring in either monthly or daily units and each observation as occuring at the midpoint of its observation interval. The estimates of these models show a statistically insignificant effect of unemployment on approval in models without a dummy variable for the Bush administration but a statistically significant effect in models with the dummy variable.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

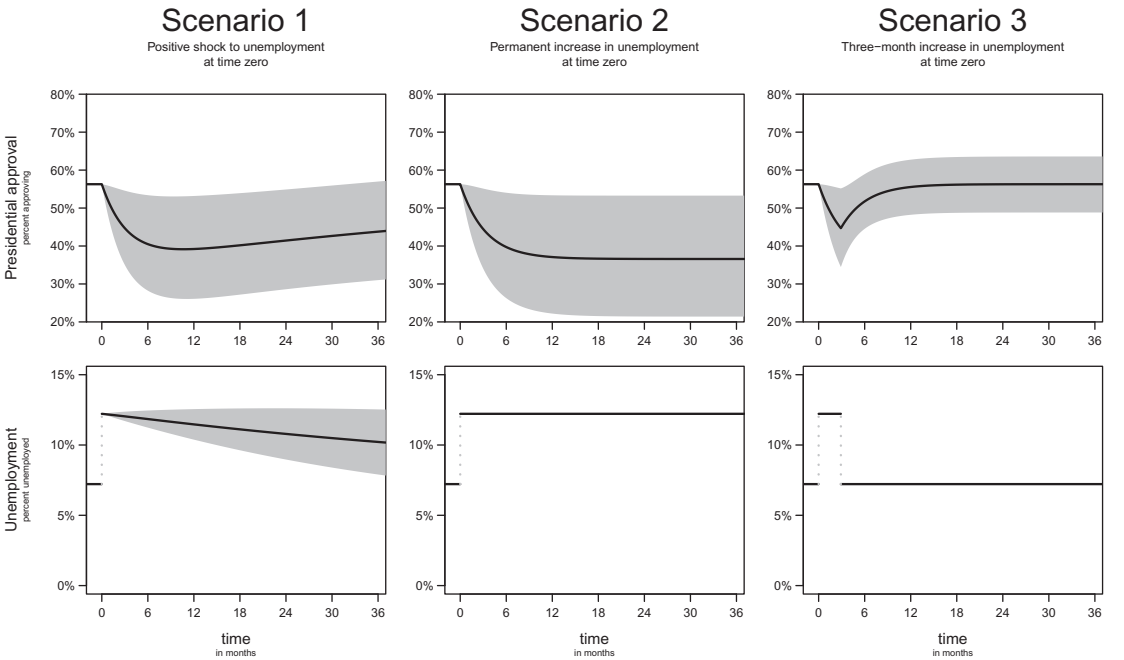


Fig. 3 Simulated scenarios under estimated relationship between unemployment and presidential approval. Shaded areas are 95% confidence regions.

these results demonstrate that MONOCAR can lead to notably different conclusions on real data sets.

5.2.3 Interpretation

Model-implied simulations can aid in the interpretation of MONOCAR models. As an example, this section will use simulations under three different scenarios to illustrate the implied effect of changes in unemployment on presidential approval based on the estimates of the first MONOCAR model in Table 2. In each case, both series begin at an equilibrium level and then show the expected change in approval implied by the model after unemployment in some way. The stochastic component of the model is omitted. Plots of these scenarios are shown in Fig. 3.

The first scenario involves a sudden increase in unemployment at time zero, but does not specify what happens after that point in time. The model implies that presidential approval continues to fall while unemployment slowly decreases for approximately nine months, after which presidential approval reaches its minimum, following the gradual erosion of unemployment. Thus, when unemployment increases suddenly, approval falls, but, as unemployment slowly declines after the shock, presidential approval eventually drops from low unemployment and then remains low while unemployment remains high.

The second scenario demonstrates the predicted effect of a long-term change in unemployment. In this scenario, unemployment suddenly increases and remains at a permanently elevated level. The predictions suggest that although presidential approval begins to fall immediately following an increase in unemployment, it may take roughly a year before approval is near its new equilibrium low following a permanent increase in unemployment.

The third and final scenario is more informative about the speed with which presidential approval is predicted to respond to a rise in unemployment. In this scenario, unemployment suddenly becomes high for three months, but then returns to its original level at the end of that period. Presidential approval is predicted to rise while unemployment is high, but declines after it returns to its original level. This scenario demonstrates the speed with which presidential approval

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

declines. Presidential approval has returned roughly half way to its original level three months after unemployment drops, which is a faster response than the model expects from unemployment.

6 Conclusion

Time series data often include complications that do not fit the assumptions of standard models like VAR. Data often include irregularly spaced observations collected over nontrivial and even overlapping intervals of time with measurement error. Opinion polls and events data provide two notable examples, but these problems are also exhibited in a variety of other types of data.

This article introduces a model, referred to as the multivariate overlapping noisy observation continuous-time autoregressive model or MONOCAR model, to address these difficult properties. By treating the data as driven by a continuous-time latent process with observations that can be noisy, noninstantaneous, and overlapping, this model captures these features and removes several potential sources of bias. The MONOCAR model also eliminates the loss of data that occurs during temporal aggregation, which may allow us to detect smaller and shorter-term effects than would otherwise be possible.

As a solution to the problem of temporal aggregation, previous work has suggested looking for a natural unit of time or conducting robustness checks across alternative levels of aggregation (Freeman 1989; Shellman 2004). Although this is undoubtedly good advice if forced to choose a time unit, MONOCAR provides an alternative: not choosing a unit. Because MONOCAR is based around a continuous-time model, there is no choice of aggregation level to be made. In some cases, the underlying processes we are modeling may truly be a discrete process and have a natural unit of time. However, if the processes are best thought of as occurring in continuous-time, modeling them as such and avoiding the choice of aggregation level may provide the best solution.

The MONOCAR model can be applied to a wide variety of research questions and data. The advantages of MONOCAR are not only theoretical but can be seen in real-world applications, where MONOCAR can lead to different conclusions than more common approaches. The development of an R library, which allows for easy estimation and inference in a maximum-likelihood framework, may facilitate the use of this model on other research projects.

Like any model, MONOCAR is not always the most appropriate tool. A simpler model, such as a VAR model, may be preferable when the data contain few of these features, whereas other models may be necessary to address other possible complications not discussed in this article. Nonetheless, MONOCAR can be a useful approach to addressing the common complications discussed in this article.

Funding

Support was provided by the Wisconsin Alumni Research Foundation through the University of Wisconsin—Madison Graduate School Fall Competition (PRJ69ZE).

References

Armstrong II, D. A. 2009. Measuring the democracy–repression nexus. *Electoral Studies* 28(3):403–12.

Beck, N. 1989. Estimating dynamic models using Kalman filtering. *Political Analysis* 1(1):121–56.

Berardo, R., and J. T. Scholz. 2010. Self-organizing policy networks: Risk, partner selection, and cooperation in estuaries. *American Journal of Political Science* 54(3):632–49.

Berlemann, M., and S. Enkelmann 2012. The economic determinants of U.S. presidential approval—A survey. SSRN Scholarly Paper ID 2029449, Social Science Research Network, Rochester, NY.

Bond, J. R., R. Fleisher, and B. D. Wood. 2003. The marginal and time-varying effect of public approval on presidential success in Congress. *Journal of Politics* 65(1):92–110.

Box-Steffensmeier, J. M., and B. S. Jones. 1997. Time is of the essence: Event history models in political science. *American Journal of Political Science* 41(4):1414–61.

———. 2004. *Event History Modeling: A Guide for Social Scientists*. Cambridge: Cambridge University Press.

Brandt, P. T., and J. R. Freeman. 2006. Advances in Bayesian time series modeling and the study of politics: Theory testing, forecasting, and policy analysis. *Political Analysis* 14(1):1–36.

———. 2009. Modeling macro-political dynamics. *Political Analysis* 17(2):113–42.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

Brandt, P. T., and J. T. Williams. 2001. A linear Poisson autoregressive model: The Poisson AR(p) model. *Political Analysis* 9(2):164–84.
Brandt, P. T., J. T. Williams, B. O. Fordham, and B. Pollins. 2000. Dynamic modeling for persistent event-count time series. *American Journal of Political Science* 44(4):823.
Brockwell, P. J. 2001. Continuous-time ARMA processes. In *Stochastic Processes: Theory and Methods*, Volume 19 of *Handbook of Statistics*, eds. D. N. Shanbhag and C. R. Rao, 149–276. Amsterdam: North Holland.
Burden, B. C., and A. Mughan. 2003. The international economy and presidential approval. *Public Opinion Quarterly* 67(4):555–78.
Carlin, R. E., G. J. Love, and C. Martínez-Gallardo. Forthcoming. Cushioning the fall: Scandals, economic conditions, and executive approval. *Political Behavior*. doi:10.1007/s11109-014-9267-3.
Carpenter, D. 2003. Why do bureaucrats delay? In In *Politics, Policy, and Organizations: Frontiers in the Study of Bureaucracy 2003*, eds. G. A. Krause and K. J. Meier, 23–40. Ann Arbor: University of Michigan Press.
Commandeur, J. J., and S. J. Koopman. 2007. *An Introduction to State Space Time Series Analysis*. Oxford: Oxford University Press.
De Boef, S., and L. Keele. 2008. Taking time seriously. *American Journal of Political Science* 52(1):184–200.
Durbin, J., and S. J. Koopman. 2012. *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.
Enns, P. K. 2014. The public's increasing punitiveness and its influence on mass incarceration in the United States. *American Journal of Political Science* 58:857–72.
Fearon, J. D. 1994. Domestic political audiences and the escalation of international disputes. *American Political Science Review* 88(03):577–92.
———. 1998. Bargaining, enforcement, and international cooperation. *International Organization* 52(02):269–305.
Freeman, J. R. 1989. Systematic sampling, temporal aggregation, and the study of political relationships. *Political Analysis* 1(1):61–98.
Freeman, J. R., J. T. Williams, and T. Lin. 1989, November. Vector autoregression and the study of politics. *American Journal of Political Science* 33(4):842–77.
Fuller, W. A. 1976. *Introduction to Statistical Time Series*. New York: John Wiley & Sons.
Goldstein, J. S. 1992. A conflict-cooperation scale for WEIS events data. *Journal of Conflict Resolution* 36(2):369–85.
Goldstein, J. S., and J. C. Pevehouse. 1997. Reciprocity, bullying, and international cooperation: Time series analysis of the Bosnia conflict. *American Political Science Review* 91(3):515–29.
Green, D. P., A. S. Gerber, and S. L. d. Boef. 1999. Tracking opinion over time: A method for reducing sampling error. *Public Opinion Quarterly* 63(2):178–92.
Gronke, P., and J. Brehm. 2002. History, heterogeneity, and presidential approval: A modified ARCH approach. *Electoral Studies* 21(3):425–52.
Hänggi, P., and F. Marchesoni. 2005. Introduction: 100 years of Brownian motion. *Chaos* 15(2):026101.
Hibbs, Jr., D. A. 1973. Problems of statistical estimation and causal inference in time-series regression models. In *Sociological Methodology*, Vol. 5, ed. D. R. Heise, 252–308. San Francisco: Jossey-Bass.
Jackman, S. 2005. December. Pooling the polls over an election campaign. *Australian Journal of Political Science* 40:499–517.
Jones, R. 1984. Fitting multivariate models to unequally spaced data. In *Time Series Analysis of Irregularly Observed Data*. vol. 25 of *Lecture Notes in Statistics*, ed. E. Parzen, 158–188. New York: Springer.
Jones, R., and P. Tryon. 1987. Continuous time series models for unequally spaced data applied to modeling atomic clocks. *SIAM Journal on Scientific and Statistical Computing* 8(1):71–81.
Jones, R. H. 1981. Fitting a continuous time autoregression to discrete data. In *Applied Time Series Analysis II*, ed. D. F. Findley, 651–682. New York: Academic Press.
———. 1985. Time series analysis with unequally spaced data. In *Time Series in the Time Domain*. vol. 5 of *Handbook of Statistics*, eds. E. J. Hannan, P. R. Krishnaiah, and M. M. Rao, 157–178. Amsterdam: North Holland.
Jones, R. H., and L. M. Ackerson. 1990. Serial correlation in unequally spaced longitudinal data. *Biometrika* 77(4):721–31.
Jones, R. H., and F. Boadi-Boateng. 1991. Unequally spaced longitudinal data with AR(1) serial correlation. *Biometrics* 47(1):161–75.
Kellstedt, P., E. McAvoy, and J. A. Stimson. 1993. Dynamic analysis with latent constructs. *Political Analysis* 5(1):113–50.
Kernell, S. 1978. Explaining presidential popularity. *American Political Science Review* 72(2):506.
Kriner, D. L. 2006. Examining variance in presidential approval: The case of FDR in World War II. *Public Opinion Quarterly* 70(1):23–47.
Manger, M. S., M. A. Pickup, and T. A. B. Snijders. 2012. A hierarchy of preferences: A longitudinal network analysis approach to PTA formation. *Journal of Conflict Resolution* 56(5):853–78.
Martin, A. D., and K. M. Quinn. 2002. Dynamic ideal point estimation via Markov chain Monte Carlo for the U.S. Supreme Court, 1953–1999. *Political Analysis* 10(2):134–53.
McAvoy, G. E. 1998. Partisan probing and democratic decisionmaking rethinking the NIMBY syndrome. *Policy Studies Journal* 26(2):274–92.
Mueller, J. E. 1970. Presidential popularity from Truman to Johnson. *American Political Science Review* 64(1):18.
Pickup, M. A., and R. Johnston. 2007. Campaign trial heats as electoral information. *Electoral Studies* 26(2):460–76.
———. 2008. Campaign trial heats as election forecasts. *International Journal of Forecasting* 24(2):272–84.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press

Pickup, M. A., and C. Wlezien. 2009. On filtering longitudinal public opinion data. *Electoral Studies* 28(3):354–67.
Roper Center. 2013. *iPoll databank*. Storrs, CT.: The Roper Center, University of Connecticut.
Schweppe, F. C. 1965. Evaluation of likelihood functions for Gaussian signals. *IEEE Transactions on Information Theory* 11(1):61–70.
Shellman, S. M. 2004. Time series intervals and statistical inference: The effects of temporal aggregation on event data analysis. *Political Analysis* 12(1):97–104.
Stimson, J. A. 1999. *Public Opinion in America: Moods, Cycles, and Swings*. Boulder: Westview Press.
———. 2008. Using WCALC and the dyad ratios algorithm. http://www.unc.edu/~jstimson/Software_files/wcalc.pdf (accessed January 22, 2015).
Streb, J. M., D. Lema, and P. Garofalo. 2012. Fall. Temporal aggregation in political budget cycles. *Economía* 13(1):39–78.
Tiao, G. C., and W. W. Wei. 1976. Effect of temporal aggregation on the dynamic relationship of two time series variables. *Biometrika* 63(3):513–23.

https://doi.org/10.1093/pan/mpu020 Published online by Cambridge University Press