Psychological Methods

A Tutorial and Methodological Review of Linear Time Series Models: Using R and SPSS

Jesús F. Rosel, Sara Puchol, Marcel Elipe, Patricia Flor, Francisco H. Machancoses, and Juan J. Canales

Online First Publication, November 13, 2025. https://dx.doi.org/10.1037/met0000794

CITATION

Rosel, J. F., Puchol, S., Elipe, M., Flor, P., Machancoses, F. H., & Canales, J. J. (2025). A tutorial and methodological review of linear time series models: Using R and SPSS. *Psychological Methods*. Advance online publication. https://dx.doi.org/10.1037/met0000794

AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Methods

© 2025 The Author(s)
ISSN: 1082-989X

https://doi.org/10.1037/met0000794

A Tutorial and Methodological Review of Linear Time Series Models: Using R and SPSS

Jesús F. Rosel[1], Sara Puchol[2], Marcel Elipe[1], Patricia Flor[3], Francisco H. Machancoses[1], and Juan J. Canales[4]

[1] Faculty of Health Sciences, Universitat Jaume I

[2] Independent Researcher, Castellón, Spain

[3] Faculty of Health Sciences, Valencia International University

[4] Faculty of Health and Environmental Sciences, Auckland University of Technology

Abstract

This article introduces autoregressive (AR) linear models to psychology students and researchers through a step-by-step approach using SPSS and R. Despite their relevance, AR models remain underutilized in behavioral sciences, possibly due to conceptual challenges and difficulties interpreting autocorrelation and seasonality. Our aim is to simplify their implementation by presenting time series models as special cases of linear regression, using accessible language and practical examples. The article illustrates AR estimation using real data, incorporating lagged values as predictors of the dependent variable. Residual diagnostics, a frequently overlooked aspect in applied research, receive special attention, including figures and statistical tests. As Kmenta (1971) demonstrated, serially correlated residuals can lead to artificially low *p* values for the parameter estimates, potentially resulting in explanatory variables being deemed significant when they truly are not. To promote understanding, we offer intuitive visualizations and clear decision rules for model building, lag selection, and seasonality detection. We compare polynomial and AR models using the confounding test. The data set and annotated R and SPSS scripts are included to support replication and help readers learn basic syntax. We also discuss conceptual and practical limitations of moving average, integration (I), and exponential smoothing models, emphasizing the practical advantages of AR-only models in psychological contexts. Throughout, we stress the importance of aligning statistical models with theoretical assumptions and the temporal structure of data. By combining step-by-step explanations, visual guidance, and real-data applications, this tutorial provides a practical foundation for incorporating AR models into applied psychological research.

Translational Abstract

This tutorial introduces the analysis of univariate time series (TS) using autoregressive (AR) linear models with SPSS and R software. It illustrates how AR models can be estimated as ordinary linear regressions, eliminating the need for specialized TS software. When residuals meet the assumptions of white noise (normality, homoscedasticity, and absence of autocorrelation), results from ordinary regression and classical TS estimation converge. We highlight the advantages of AR models in psychological research over alternative TS approaches, especially in contexts where model transparency and flexibility are valued. Key modeling considerations are discussed, such as assessing the inclusion of explanatory variables that seem unnecessary but are retained due to prior hypotheses (overparameterization), and the need to detect seasonality—systematic patterns at fixed intervals, such as weekly or annual cycles. The tutorial also warns of frequent

Emilio Ferrer served as action editor.

Jesús F. Rosel https://orcid.org/0000-0002-0920-032X

Juan J. Canales https://orcid.org/0000-0003-2884-5227

Some of the ideas presented in this tutorial were first introduced in doctoral courses on Time Series in Psychology taught by Jesús F. Rosel. These courses were offered at the Universidad de Málaga, coordinated by A. Fierro (1985–1992); at the Universitat Jaume I of Castellón, coordinated by R. A. Clemente (1993–2002); and in a training course for researchers at the Universidad Autónoma de Madrid, coordinated by V. Ponsoda (2017). The authors declare that they have no conflicts of interest.

Open Access funding provided by Auckland University of Technology: This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0; https://creativecommons.org/licenses/by/4.0). This license permits copying and redistributing the work in any medium or format, as well as adapting the material for any purpose, even commercially.

Jesús F. Rosel served as lead for conceptualization, formal analysis, and methodology and contributed equally to writing–original draft. Sara Puchol served as lead for data curation, investigation, and software and contributed equally to visualization. Marcel Elipe contributed equally to validation and served in a supporting role for writing–review and editing. Patricia Flor served as lead for data curation and served in a supporting role for software and writing–review and editing. Francisco H. Machancoses contributed equally to formal analysis and methodology and served in a supporting role for writing–review and editing. Juan J. Canales served as lead for supervision and writing–review and editing and served in a supporting role for conceptualization and methodology. Marcel Elipe and Francisco H. Machancoses contributed equally to software.

Correspondence concerning this article should be addressed to Juan J. Canales, Faculty of Health and Environmental Sciences, Auckland University of Technology, AZ Building, 90 Akoranga Drive, Northcote, Auckland 0627, New Zealand. Email: juan.canales@aut.ac.nz

software-specific errors in parameter estimation, particularly in the intercept, which may mislead interpretation. An applied example demonstrates how AR models outperform straight-line or polynomial regression models by properly addressing autocorrelated residuals, thereby avoiding false significance (inflated Type I error rates). By learning to implement lagged variables in univariate TS models, researchers are better prepared to engage with more advanced structures such as pooled TS, panel data, or dynamic systems. This step-by-step guide aims to strengthen both conceptual understanding and technical proficiency, serving as a practical foundation for students and professionals beginning in longitudinal and TS analysis.

Keywords: autoregressive models, time series models, ordinary regression, lagged variables, R tutorial

Supplemental materials: https://doi.org/10.1037/met0000794.supp

This tutorial offers a practical approach to analyzing univariate autoregressive (AR) time series (TS) data, guiding you through each step of longitudinal data analysis to deepen your understanding of time-related regression processes. We will also show the reader how to ensure that residuals from these analyses are "white noise," meaning they are free from significant correlations.

With the advancement of new automated methods for data collection in both laboratory and real-world settings, a significant amount of temporal data is being generated (Hamaker & Wichers, 2017). This has led to various methodological labels depending on the research focus. Indeed, numerous studies are categorized under the umbrella of longitudinal research, using terms such as experience sampling method, ecological momentary assessment, real-time data collection, intensive longitudinal design, and panel data design.

However, when it comes to data analysis, there appears to be a disproportionate lack of studies utilizing TS methods (Ariens et al., 2020; Smith, 2012). This disparity can be attributed to two key factors. First, TS analysis is rarely included in doctoral curricula, appearing primarily in programs focused on quantitative research, where it is often limited to just a few credits. Second, it is frequently regarded as a complex and highly specialized subject, which may discourage wider adoption. However, in the following pages, we will show that AR TS models are, in fact, a special case of regression analysis; drawing on hypothesis from "dynamic modeling" or "general-to-specific" approach in statistical modeling (Campos et al., 2005; Florens & Mouchart, 1985; Hendry, 2024; Hendry et al., 1984; Huckfeldt et al., 1982; Pesaran, 2016), we aim to show that these models can be both accessible and practically applicable.

The Importance of TS Analysis in Psychological Research

Historically, TS analysis was applied in behavioral tracking (e.g., Skinner, 1938) and formal single-case designs (e.g., Shapiro, 1961). With the emergence of daily diary methods and ecological momentary assessment (Bolger et al., 2003), TS analysis has become central to modeling dynamic changes in mood, cognition, and behavior in real-world contexts.

TS models can be classified by their domain and by the number of subjects and independent variables (IVs) involved. Time-domain models assess temporal dependencies through their structural formulation (e.g., autoregressive [AR], differencing/Integration [I], moving average [MA], and the general synthesis of the three, ARIMA), whereas frequency-domain models focus on identifying cyclical patterns (Shumway & Stoffer, 2017). Unlike time-domain approaches, which model how a value at time t (Y_t) depends on past values, frequency-domain models decompose the series into a sum of sine and cosine waves representing different cycles. However, cyclicity can also be captured within time-domain models through the inclusion of seasonal components, which are conceptually equivalent to cyclical variation in the frequency domain. This tutorial focuses on linear time-domain models, as they are more accessible and widely used in psychological research.

TS models can also be classified by the number of individuals and IVs included. A univariate TS model examines one variable over time for a single individual, capturing intraindividual temporal patterns. In contrast, multivariate TS models may analyze multiple IVs for a single case (common in intervention studies) or incorporate data from multiple individuals measured over time, using one or more variables (as in panel or multilevel TS designs). The selection of model structure critically shapes the insights that can be drawn about temporal dynamics in psychological research.

Series length influences the appropriateness of different TS models, so for short series (fewer than 30 observations), parameter estimates tend to be unstable, making approaches such as Bayesian methods or bootstrapping more suitable. Medium-length series (30–50 observations) and long series (more than 50) permit the estimation of more complex lag structures. Regardless of length, it is essential to conduct diagnostic checks, including the autocorrelation function (ACF), and the partial ACF (PACF), and tests for stationarity or long-run behavior, to verify key assumptions such as residual independence, homogeneity of variance, and overall model stability.

Despite the availability of advanced techniques, many researchers still rely on polynomial regression (Fried et al., 2022; Kuba & Scheibe, 2017), partly due to its familiarity. However, linear TS models are also regression-based and conceptually accessible to most applied researchers (Hendry, 2024; Voelkle & Oud, 2013). Their combination of interpretability and flexibility makes them particularly well suited as an entry point for modeling more complex psychological TS.

Linear TS Analysis

Focusing on the analysis of a univariate TS, this guide teaches how past data points can predict future values. Our aim is to simplify complex concepts, minimizing the use of equations in favor of a stronger emphasis on understanding the procedures and the ideas behind them. We use R (R Core Team, 2023a) and IBM SPSS (IBM, 2022) to illustrate these techniques, providing a side-by-side guide to enrich the reader's learning experience. This tutorial targets students and professionals who are new to longitudinal data analysis, but it might also help those seeking more advanced knowledge and additional resources (Box et al., 2016; Box-Steffensmeier et al., 2014; Shin, 2017; Vandaele, 1983).

By building univariate AR models by hand, the reader will gain insights into the AR process, including how previous observations

can forecast future ones and the importance of lags in enhancing prediction accuracy. This practice not only clarifies key statistical concepts like serial dependency and autocorrelation (AC) but also lays the groundwork for tackling more complex data designs involving multivariate TS, with multiple subjects or variables. Although manual construction of AR models requires some more effort than automated software, especially for large data sets, the conceptual advantages are invaluable, setting a strong foundation for further exploration into multivariate TS models, such as panel data or pooled TS, and even advanced analyses, like multilevel modeling or structural equation modeling (SEM) of multivariate TS.

In this research, we will present the basic properties of AR models, their advantages over other possible models, the phases of TS analysis, and the basic hypotheses of our research; this will be exemplified with some temporal data on how to carry out the analysis of these data using the R and SPSS programs.

Serial Dependence and AC

In TS analysis, the concept of serial dependence is fundamental. This principle suggests that human behavior (whether physical, psychological, consumer-related, or social) at any given time, for example, on a particular day, is influenced by past behaviors. In turn, today's actions will shape future behaviors, creating a continuous chain of influence over time. Statistically, this behavioral continuity results in AC of the observed variable, meaning that a time-dependent variable exhibits a correlation with its own past values.

When the values of a TS are lagged and correlated with the original (unlagged) series, significant correlations typically emerge, reflecting this dependency. This self-correlation of a temporal variable with its own past values is known as AC, which is essential for identifying the underlying AR model governing the temporal structure of the observed behavior. A crucial aspect of TS analysis is to model the serial dependence of the variable, while ensuring at the final stage of the process that the residuals do not exhibit AC, as they should behave as white noise to validate the model's assumptions.

The Consequences of Residual AC

A Type I error, also known as a false positive, occurs when a researcher incorrectly rejects a true null hypothesis. Type I error occurs when a statistical test incorrectly rejects a true null hypothesis, or in other words, to infer that there is an effect when really there is not. We will see that when there is autocorrelation of residuals (ACRs), there is a tendency to commit Type I errors with respect to the IVs utilized. This issue was first highlighted by several authors in the field of statistics (Aitken, 1936; Cochrane & Orcutt, 1949), who proposed the use of AR models to remove ACR.

In psychology, there has been an ongoing debate regarding the appropriateness of using AR models for TS data. However, several authors have shown that failing to apply AR models when analyzing TS data often results in significant ACR (Arnau & Bono, 2001; Hibbs, 1974; Huitema et al., 1999; Jones et al., 1977; Kazdin, 1982; Suen & Ary, 1987). Most of these findings have been derived from simulation studies, stressing the importance of proper TS modeling to avoid biased conclusions.

Kmenta (1971, pp. 274–281) demonstrated, in the context of statistical economics, that when ACR is significant, utilizing ordinary least squares (OLS) or similar methods for parameter estimation leads to underestimated error variances. Consequently, the standard errors of the parameters, included in the estimators' denominators, are similarly underestimated. This situation causes overestimation of values in bs, ts, zs, … F, R^2, … statistics, leading to inefficiency and an increased risk of Type I errors. Therefore, if ACR is not significant, using OLS or similar procedures becomes appropriate for reliable parameter estimation in temporal regression.

This finding has had a significant impact on econometrics, as nearly all textbooks include a dedicated chapter on ACR, as noted in Kmenta's (1971) research. Moreover, virtually all TS analyses published in economics and statistics journals (whether univariate, multivariate, or panel data studies) conduct residual diagnostics to assess model validity. However, this type of analysis is often overlooked in most longitudinal research within the social or health sciences, despite its critical role in ensuring accurate statistical inference.

In Kmenta's manual, a straightforward illustration is provided regarding ACR when it follows an AR(1) pattern. It begins by assuming a forecast model where residuals (e_t) are autocorrelated with a value of ρ, represented as $e_t = \rho e_{t-1} + \varepsilon_t$. This formulation leads to an evaluation of the variance of e_t:

$$\text{var}(e_t) = \frac{\text{var}(\varepsilon_t)}{1 - \rho^2}. \quad (1)$$

Here, several aspects should be noted. The first is that if $\rho = 0$, then $\text{var}(e_t) = \text{var}(\varepsilon_t)$. The second aspect is that for values of $\rho \neq 0$, $\text{var}(e_t) > \text{var}(\varepsilon_t)$. Also, the value of $\text{var}(e_t)$ will be larger the greater the absolute value of ρ, making it easier to make Type I errors; this confirms simulation studies with TS.

As a conclusion to this section on ACR, we present a fictitious example. Suppose a researcher hypothesizes that positive mood is a function of TV viewing hours, with both variables being TS data collected from the same individual over 60 days. The researcher finds a positive and significant relationship between these temporal variables and concludes that positive mood depends on the number of hours of TV watched per day, influencing the person's mood on the same day.

However, let us assume that the researcher did not analyze the residuals, and they exhibit significant ACF and PACF values. In this case, it is highly likely that failing to account for ACR has led to underestimation of the error variance, $\text{var}(\varepsilon_t)$, resulting in inflated b, t, F, … statistics. Consequently, the null hypothesis is likely to be rejected when it is actually true (Type I error).

In our example, the estimated b value does not reflect the true effect of TV viewing hours on positive mood. The actual effect can only be obtained by including an appropriate IV, in addition to TV viewing hours, that eliminates AC, leading to more accurate b values. Notice the paradox: introducing a nonsignificant IV that still allows ACR to persist increases the likelihood that the IV appears statistically significant, causing its effect to be artificially overestimated. Thus, this mistake can lead to incorrect conclusions and result in misguided recommendations for potential therapeutic interventions.

Brief Literature Review

Despite the inherently longitudinal nature of single-case designs, TS analysis remains rarely applied in psychological research. Molenaar (2004) emphasized the need for psychology to evolve as an idiographic science, advocating for models that capture intraindividual dynamics over time. Earlier, Molenaar (1985) had already

advanced this agenda by proposing a dynamic factor analysis model that integrated latent variables with temporal structure, laying conceptual groundwork for later methodological developments. However, empirical reviews by Smith (2012) and Kazdin (2019, 2021) indicate that fewer than 2% of single-case studies employ formal statistical TS models. Instead, many rely on visual inspection or polynomial regression, often without evaluating key assumptions such as autocorrelated residuals (Baek & Ferron, 2013; Shadish & Sullivan, 2011).

These studies highlight several key findings: (a) There is a growing shift from visual inspection toward inferential statistical methods (Hammond & Gast, 2010; McDonald et al., 2017; Smith, 2012; Tanious & Onghena, 2021); (b) polynomial models remain popular but frequently fail to address autocorrelated residuals (Baek & Ferron, 2013; Hyndman & Athanasopoulos, 2021; Jebb et al., 2015; Rosel et al., 2019); (c) AR structures are seldom modeled explicitly (Shadish & Sullivan, 2011); (d) SEM dominates intensive longitudinal designs research, yet often excludes AR or vector AR (VAR) components (Andersen, 2022; Bollen & Brand, 2010; Bringmann, 2021; Ferron et al., 2009; Lüdtke & Robitzsch, 2022; Nestler & Humberg, 2024; Robinaugh et al., 2020; Segerstrom et al., 2017); and (e) methodological reviews frequently note the absence of residual diagnostics and formal model validation (Hammond & Gast, 2010; McDonald et al., 2017; Tanious & Onghena, 2021).

Other issues persist, as revealed in our review of longitudinal studies: (a) Many studies lack explicit hypotheses about lag structures or seasonality, even when 7-day cycles are plausible; (b) aggregation across individuals or sessions often results in a loss of intraindividual variability and may introduce Yule-Simpson effects, where trends observed in raw data disappear or reverse when averaged across categories (Hintzman, 1980; Kievit et al., 2013; Yule, 1903). This is particularly problematic in pre–post treatment designs and crossover trials, where data are averaged by phase, obscuring important temporal dynamics (Cundiff & Matthews, 2018; Dobrow et al., 2018; Kiesner et al., 2016; Kumar et al., 2019); (c) automated modeling tools such as AUTOBOX (Reilly, 2019) are sometimes used without prior specification of seasonality in the hypotheses (Maican et al., 2016); (d) group comparisons are frequently conducted using separate models, rather than incorporating group membership as dummy variables within a single model, limiting statistical power and interpretability (Munafò et al., 2008; Murray et al., 2011); (e) incomplete reporting and interpretation remains common: many studies omit the final TS model equation and fail to interpret estimated parameters or seasonal patterns (Antoniou et al., 2021); (f) omission of lagged and seasonal predictors is widespread. TS regression models often rely solely on concurrent predictors, for example, $Y_t = f(X1_t, X2_t, X3_t, \ldots)$, ignoring lagged effects that could better capture temporal dependencies, for example, $Y_t = f(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-12}, Y_{t-24}, X1_t, X1_{t-1}, \ldots)$ (Bye, 2007); and (g) residual diagnostics are rarely performed, and even when significant autocorrelation is detected, few studies attempt to refine the model accordingly (Chowdhury & Wang, 2009).

Nonetheless, applied research is increasingly demonstrating the potential of TS models. Bringmann et al. (2013) modeled affective dynamics in individuals using AR structures, showing how emotional states influence one another over time. Fisher and Boswell (2016) applied idiographic AR models to track within-person therapeutic progress, identifying temporal markers of clinical change. Ferrer and Zhang (2009) reviewed TS applications in cognitive research, illustrating how TS methods can capture learning, memory, and attention dynamics. Together, these studies exemplify how TS models enable rigorous analysis of intraindividual change processes in psychological science.

Recent developments, such as dynamic structural equation modeling (DSEM; Asparouhov et al., 2018), dynamic factor analysis (Browne & Nesselroade, 2005; Molenaar, 1985), and psychometric network models (Borsboom et al., 2021), have integrated TS logic into broader latent variable frameworks. Although these approaches introduce greater complexity, they remain grounded in the foundational principles of univariate AR models. As advanced models such as DSEM and psychometric networks gain traction, their underlying structure continues to reflect the time-dependent dynamics first formalized in univariate TS models.

Many longitudinal studies in psychology could benefit from moving beyond polynomial trend models and adopting AR-based TS methods that explicitly account for AC, lagged effects, and temporal dependencies. Doing so would enhance causal inference, reduce model misspecification, and better align statistical modeling with the inherently dynamic nature of psychological processes.

TS Analyses in Practice

A univariate TS is a sequence of data points from a single variable, collected at equal time intervals from an individual. Its key characteristic, temporal ordering, distinguishes it from other data types, aiding in both understanding data properties and forecasting future events. Consider a time variable, Y_t (e.g., daily cigarette consumption), where Y denotes the variable's substantive content, and t indicates the measurement time, ranging from 1 to N. Thus, the variable's order is $Y_1, Y_2, Y_3, \ldots, Y_t, \ldots, Y_{N-1}, Y_N$, with subscripts 1, 2, 3, …, t, …, $N-1$, N being the day number.

A Y_t value depends on its past values ($Y_{t-1}, Y_{t-2}, Y_{t-3}$ …) and seasonally recurring values ($Y_{t-S}, Y_{t-2S}, Y_{t-3S}$ …), where S denotes the seasonality period, such as 7 days for a week. For example, in someone with a smoking habit, daily cigarette numbers can be influenced by consumption in preceding days and similar past seasonal days. Thus, the cigarettes smoked on a specific day, say December 16, Saturday, relate to those smoked on preceding immediate days (Friday 15, Thursday 14, etc.) and on past Saturdays (December 9, December 2, etc.). A TS analyst's role involves identifying these significant immediate and seasonal lags (e.g., $t-1$, $t-2$, $t-3$ … and $t-7$, $t-14$, $t-21$ …) in a comprehensive model, $Y_t = f(Y_{t-1}, Y_{t-2}, Y_{t-3}\ldots, Y_{t-7}, Y_{t-14}, Y_{t-21}\ldots)$, and determining the coefficients' values in this temporal regression.

In AR TS models within psychology, the sign of AR coefficients reveals how past psychological states influence present behaviors. A positive AR coefficient suggests continuity, so if stress is high one day, it is likely to remain elevated the next. Conversely, a negative AR coefficient indicates an inverse relationship. For example, in eating regulation, a day of excessive consumption may lead to compensatory restrictions the following day, and vice versa (Velicer et al., 1992).

The appropriateness of AR coefficient signs depends on the psychological process studied. Stable behaviors typically show positive AR coefficients, reflecting persistence and complacency. In contrast, individuals struggling with behavioral control often exhibit negative AR coefficients, as compensatory mechanisms create an alternating pattern. This is evident in daily binge eating cycles, where overeating a day is followed by strict control the next day, only to relapse again,

forming a "sawtooth-like" trajectory of fluctuation above and below the average.

This article underscores that TS models are essentially specialized regression cases, eliminating the need for dedicated TS software. Any standard linear regression tool can suffice, provided the original Y_t values are appropriately lagged. For instance, the simplest TS model, AR(1), follows:

$$Y_t = b_0 + b_1 Y_{t-1} + e_t, \tag{2}$$

where b_0 is the intercept and b_1 is the AR coefficient for the lagged term Y_{t-1}. This corresponds to the Box and Jenkins (1970) formulation, where these parameters are denoted as ϕ_0 and ϕ_1, respectively. The error term e_t is analogous to ε_t in the Box and Jenkins model framework. Here, e_t (or ε_t) is a "white noise" process, or random, uncorrelated at different times, with a mean of zero and constant variance, ensuring unpredictability in the e_t series.

We introduce the concept of stationarity, a condition where a process's statistical properties, like mean and variance, remain constant over time. It is crucial for the original series Y_t to exhibit stationarity for accurate analysis and forecasting. Regarding AC, each empirical series Y_t has its own AC process, given by the simple and partial correlations between the values of the series and their corresponding lags; but e_t, like white noise, must have zero-value ACs.

George U. Yule, in collaboration with Gilbert Walker, made significant contributions to the development of AR models in the early 20th century; Yule (1927) introduced the method of AR vaiables in his work on TS analysis, and Walker (1931) further developed these ideas. Peter Whittle, in his 1951 thesis, made significant contributions to the MA component. Box and Jenkins (1970) developed a systematic methodology for TS analysis, encompassing AR, MA, and integrated (I) components, which they extended to incorporate seasonal patterns, resulting in the general seasonal ARIMA model, denoted SARIMA(p,d,q)(P,D,Q)S procedure (Velicer & Molenaar, 2013).

To distinguish between AR and MA models, Box and Jenkins (1970) demonstrated that an AR process of order p, AR(p), is characterized by an ACF that declines gradually, often exponentially, while the PACF exhibits a sharp cutoff after lag p (only the first p lags are significantly different from zero). In contrast, an MA process of order q, MA(q), displays the opposite pattern: the ACF cuts off abruptly after lag q, with significant values only at the first q lags, whereas the PACF decays gradually, often showing significance beyond lag q.

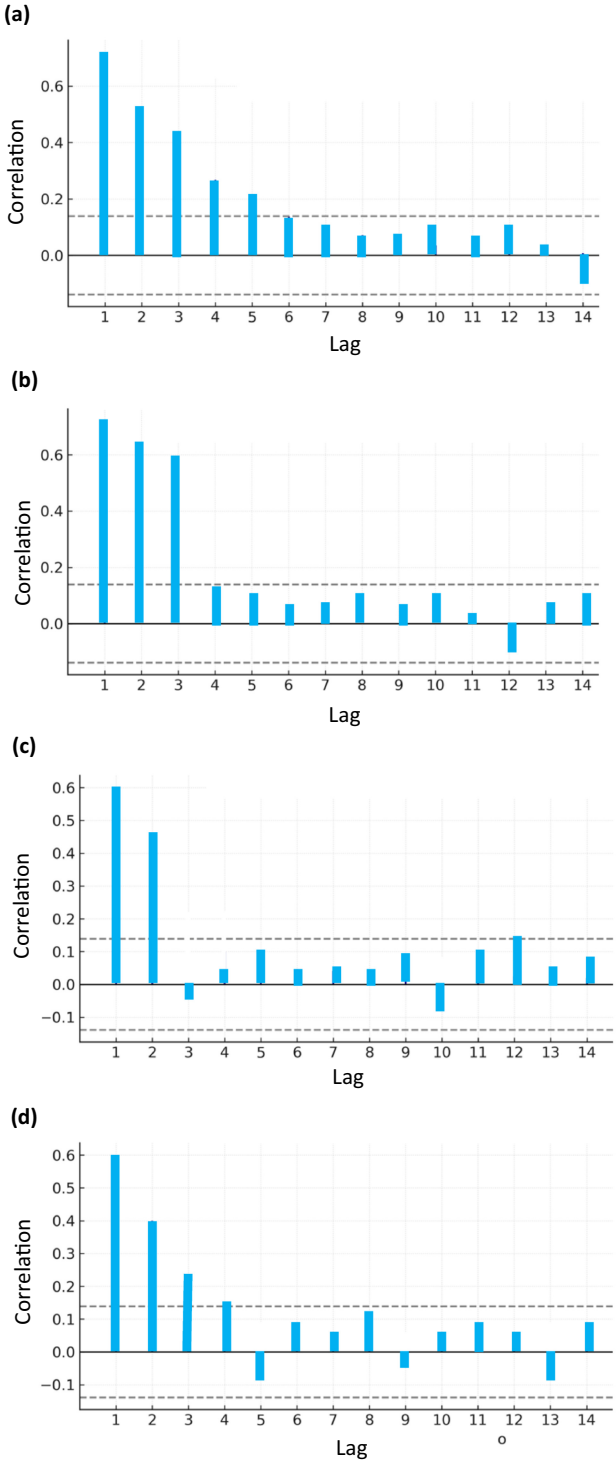
In an AR(3) model, the current value of the series depends on the preceding three observations. The ACF displays a smooth, gradual decline across lags, while the PACF shows significant spikes at lags 1, 2, and 3 before dropping to near-zero values of PACF, indicating that only the first three lags contribute directly once earlier lags are accounted for.

In contrast, an MA(2) model depends on recent random shocks, that is, on unpredictable changes or innovations (e_t) that occurred in previous time points. In this case, the ACF shows a sharp cutoff after lag 2, while the PACF decays gradually, reflecting the indirect effect of past shocks. Figure 1 displays these typical ACF and PACF patterns for AR(3) and MA(2) models.

Seasonal versions of these models follow the same logic but at periodic lags. For example, a seasonal AR(3) with Period 7, AR(3)$_7$, will show ACF decay at seasonal lags (7, 14, 21, …),

Figure 1

Simulated Autocorrelation and Partial Autocorrelation Functions of AR(3) and MA(2) Models



Note. Panel (a) shows the ACF of the AR(3) model; Panel (b) shows the PACF of the AR(3) model; Panel (c) shows the ACF of the MA(2) model; and Panel (d) shows the PACF of the MA(2) model. AR = autoregressive; MA = moving average; ACF = autocorrelation function; PACF = partial ACF. See the online article for the color version of this figure.

while its PACF cuts off after lag 21. Conversely, a seasonal $MA(2)_7$ shows ACF spikes at lags 7 and 14, followed by a cutoff, while the PACF decays slowly at seasonal lags. Figure 2 illustrates these seasonal AC structures.

We focus exclusively on AR(p) and AR(P)$_S$ models in this tutorial, arguing their suitability over Box and Jenkins models in psychological contexts. Sometimes researchers use *I* models when they observe a possible trend in the series, but the *I* models increase or decrease indefinitely; however, in psychology, the variables have natural bounds (Goldstein, 2019; Jonas & Stollberg, 2021). Differentiating a TS could obscure its inherent statistical properties and, from a methodological standpoint, leaves the long-term trend (LTT) indeterminate (Engle & Granger, 1987), see Appendix A.

MA models are generally discouraged in psychology because they often rely on data mining without solid theoretical foundations, since models in psychology should be grounded in hypotheses (Sternberg, 2019); additionally, simulation studies have shown that distinguishing between AR and MA models in exploratory ACF and PACF can be challenging (Campos et al., 2005; Hendry & Trivedi, 1972). Another argument for discarding MA models is that, according to the invertibility theorem by Box and Jenkins (1970), an MA(q) model can be represented as an AR(∞) model, or MA(q) $\Rightarrow$ AR(∞). However, in practice, an MA(q) model tends to fit well when it is reexpressed as an AR model with one or two more components than q, or MA(q) $\Rightarrow$ AR($q+1$, or $q+2$) (Rosel & Elósegui, 1994). The exponential smoothing models are typically reducible to *ARIMA*(0,1,1) models (Hyndman & Athanasopoulos, 2021), aligning them with the limitations associated with both Box and Jenkins *I* and *MA* approaches.

We believe that Box and Jenkins *I* models would only be justified in special cases where there is a temporal interval during which the variable level increases or decreases, necessitating a "piecewise statistical fit." However, these changes should be explained by their corresponding external or internal biological, psychological, or social causes (e.g., illness, development, intervention, etc.).

In psychological research, formulating and testing hypotheses grounded in content theories is vital for insightful, predictive data analysis (Freedman, 2012a; Sternberg, 2019; Wit et al., 2012). These steps form the core of our scientific method, aiding in the exploration and validation of theoretical concepts. TS model analysis typically unfolds in phases (Box et al., 2016; Shumway & Stoffer, 2017): (a) identification (exploratory data analysis, preliminary model evaluation …), (b) estimation (parameter fitting and preliminary results), (c) diagnostic checking (model evaluation), (d) refinement (model adjustment) and validation, and (e) interpretation of the final model from both statistical and substantive perspectives. It is recommended to analyze the parameters obtained in the final model, including both simple lag coefficients (b_1, b_2, b_3, …) for the IVs Y_{t-1}, Y_{t-2}, Y_{t-3}…, as well as seasonal lag coefficients (b_S, b_{2S}, b_{3S}, …) for the IVs Y_{t-S}, Y_{t-2S}, Y_{t-3S}, … In the empirical example analyzed in this study, a substantive interpretation of the obtained results will be provided. Each phase is iterative, often requiring back-and-forth adjustments to perfect the model.

In phase (c) of TS analysis, diagnostic checking, analyzing residuals is crucial. Under the assumptions of normality, homoskedasticity, and absence of ACR (Freedman, 2012b; Kmenta, 1971), the OLS estimator remains the best linear unbiased estimator (BLUE). However, to prevent biased standard errors, it is essential to verify the absence of ACR.

If the residuals of a TS model are not white noise, meaning they exhibit heteroscedasticity, nonnormality, or ACR, several corrective actions should be considered to improve model validity and robustness. (a) If residuals exhibit heteroscedasticity (nonconstant variance), it is advisable to apply variance-stabilizing transformations, such as logarithmic or Box–Cox transformations, to normalize variance; additionally, generalized least squares can be used to correct for unequal variance across time points. (b) If residuals deviate from normality, applying data transformations (e.g., log, square root, or inverse) can help normalize the distribution. Additionally, bootstrapping methods can be used to obtain more robust standard errors and confidence intervals. In cases where the assumption of normality is violated in small samples, Bayesian TS modeling offers an alternative, as it is less sensitive to normality assumptions. (c) If residuals exhibit AC, the first step is to check for model specification errors, revising ACF and PACF of the model residuals, and ensure that relevant lags are included. If residuals still exhibit ACR after model adjustments, Newey and West (1987) standard errors can be employed to provide robust inference. Addressing these issues ensures the validity of statistical inference, reduces the risk of Type I and Type II errors, and improves the predictive performance of TS models.

A valid TS model must satisfy certain statistical criteria to ensure its reliability for analysis and forecasting. Firstly, the residuals, representing the discrepancies between observed and predicted values, should exhibit no AC, indicating that the model has effectively captured all temporal dependencies within the data. Additionally, these residuals must be homoscedastic, and they should follow a normal distribution to facilitate valid statistical inference. Beyond the previous residual diagnostics, the model should reflect a stable LTT equilibrium among the variables, suggesting that any short-term fluctuations converge toward a consistent trend over time. This characteristic is often assessed through an LTT analysis, which tests for an LTT relationship between the dependent variable (DV) and IVs TS. Finally, it is imperative that the overall model demonstrates statistical significance, indicating that the set of IVs collectively has a meaningful relationship with the DV. Additionally, each individual coefficient should be statistically significant (except for theoretical or hypothetical conditions), underscoring the specific impact of each predictor on the DV. Adhering to these principles ensures the development of a TS model that is valid for analytical and predictive purposes.

In this tutorial, data will be analyzed using TS models with the R package ("Data Analysis 1") and with the SPSS program ("Data Analysis 2"). Additionally, the same data will be used to exemplify how an incorrect straight line analysis ("Data Analysis 3") can lead to Type I errors, assuming that the results are correct when they are not.

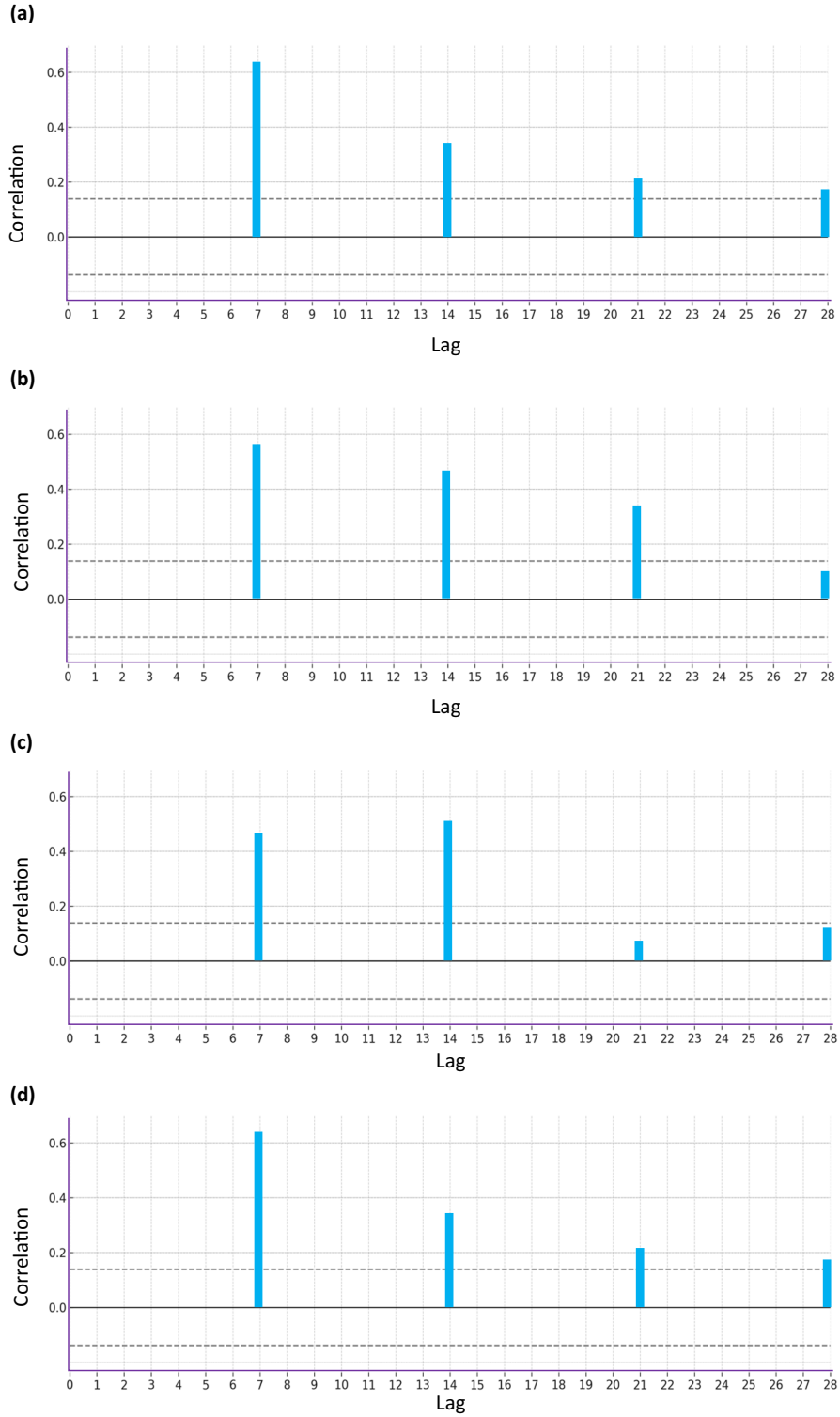
Some Methodological Considerations

One important aspect to consider in TS analysis is the temporal memory of lagged variables in an AR model. If one assumes, as a theoretical hypothesis, that a lagged IV exerts greater influence on the DV the closer it is in time (Rosel et al., 2020), then it is reasonable to impose constraints on the AR coefficients to reflect this decay. Specifically, in the AR(p)(P)$_S$ model:

$$Y_t = b_0 + b_1 Y_{t-1} + b_2 Y_{t-2} + \cdots + b_p Y_{t-p} + b_S Y_{t-S} + b_{2S} Y_{t-2S} + \cdots + b_{PS} Y_{t-PS} + e_t, \quad (3)$$

Figure 2

Simulated Autocorrelation and Partial Autocorrelation Functions of AR(3)$_7$ and MA(2)$_7$ Models (Seasonal Lag of 7 Days)



Note. Panel (a) shows the ACF of an AR(3)$_7$ model; Panel (b) shows the PACF of an AR(3)$_7$ model; Panel (c) shows theACF of a MA(2)$_7$ model; and Panel (d) shows the PACF of a MA(2)$_7$ model. AR = autoregressive; MA = moving average; ACF = autocorrelation function; PACF = partial ACF. See the online article for the color version of this figure.

one may impose:

$$b_1 \geq b_2 \geq b_3 \geq \cdots \geq b_p. \tag{4a}$$

For the immediate lags, and similarly for seasonal lags:

$$b_S \geq b_{2S} \geq \cdots \geq b_{pS}. \tag{4b}$$

Missing values are especially problematic in AR models, where each data point depends on previous ones. Deleting cases can break the time structure and reduce statistical power. A better option is maximum likelihood (ML) estimation under missing at random (MAR) assumptions, often via the expectation-maximization algorithm. Another practical solution involves state-space models, which express the TS using hidden (latent) states and observed data. These models handle missing values flexibly through algorithms like the Kalman filter (Schafer & Graham, 2002). Overall, it is important to use methods that preserve the time-order structure rather than rely on cross-sectional imputation.

TS analysis assumes that measurements are taken at equidistant intervals. When this condition is not met, a useful strategy is to introduce an additional IV that captures the timing of each observation, thus converting the model into a multivariate one. For example, if participants are instructed to complete a questionnaire once per day but do so at varying times, it is advisable to record the time of completion and include it as a categorical predictor. Another relevant variable could be the time interval since the previous measurement. Including such information helps preserve the temporal meaning of the data. Alternatively, more advanced approaches such as continuous-time modeling (e.g., available packages like `ctsem` in `R`) or the explicit use of time stamps can account for irregular intervals without requiring equidistant measurements, particularly in intensive longitudinal designs.

Hypotheses

In the description of the TS model, we have already hinted at our hypotheses; here, however, we present them more specifically. We propose that daily human behavior is subject to regular patterns based on actions taken on previous days, such that behavior on any given day is a function of behavior on immediately preceding days (Y_{t-1}, Y_{t-2}, Y_{t-3} …), showing serial dependence as observed through the correlation of the lagged variable (Hoeppner et al., 2007; Hohm et al., 2024; Markey & Markey, 2013; Meyer et al., 2016; Velicer et al., 1992). Additionally, daily human behavior exhibits a 7-day seasonality (Flor-Arasil et al., 2021; Rosel et al., 2020; Sorokin, 1964), meaning that behavior on any given day of the week, for example, a Wednesday, is influenced by behavior on preceding Wednesdays (Y_{t-S}, Y_{t-2S}, Y_{t-3S}…, or Y_{t-7}, Y_{t-14}, Y_{t-21} …). Given that the individual in our study has maintained a smoking habit for over 2 years without any intention to change, we hypothesize that the coefficients derived from the lagged IVs will be positive. This expectation aligns with findings from longitudinal studies, which have demonstrated that both nicotine dependence and average smoking levels remain stable over time, exhibiting strong positive relationships.

Empirical Data Analysis

To illustrate these concepts, we provide a case study with data, syntax (R and SPSS), and results available at https://repositori.uji.es/items/4d6bae12-a5b8-4123-a86f-20606d5d9d31 in the online supplemental materials.

Data File and Variables

We utilize data from a study on smoking behavior (Rosel et al., 2020), focusing on a 26-year-old male university student who works part-time. This data and example adhere to ethical standards, ensuring informed participant consent. The variable "P48" represents a TS detailing the daily number of cigarettes smoked by this individual over a span of 83 consecutive days, starting on a Tuesday and concluding on a Sunday, $\overline{Y_t} = 23.19$, $SD(Y_t) = 3.50$.

Data Analysis 1. TS With the `R` Package

To prepare for our analysis and graphical representation in R, we first convert the data from Excel to an R-compatible format using R Script 1 (R Core Team, 2023a). All scripts with brief explanations are available in the "R Scripts.docx" file included in the online supplemental materials. An important aspect for beginners: the "P48" variable data are in vertical orientation. The data sets "P48.rds" and "P48.xlsx" include three variables: "ID_Day": day number from 1 to 83, "Week_Day": day of the week, coded as 1 = *Monday* through 7 = *Sunday*, and "P48": the daily count of cigarettes smoked.

Figure 3 displays the trajectory of daily cigarette consumption. The vertical lines in the figure highlight each Sunday's data point, illustrating weekly patterns. The syntax for generating this graph is detailed in R Script 2a (Nordmann et al., 2022; Wickham, 2016).

Descriptive Statistics

The average (M) daily cigarette consumption "P48" is 23.19, with a standard deviation of 3.50. The number of cigarettes smoked ranges from a minimum of 17 to a maximum of 31. To further explore potential weekly trends, Script 2b guides the creation of Figure 4, which plots the mean daily consumption and its variation (± 1.96 *SD*) across different days of the week. Notably, Saturdays exhibit the highest average consumption ($M = 28.33$), whereas Mondays show the lowest ($M = 20.45$), suggesting a weekly cyclic pattern.

Identification

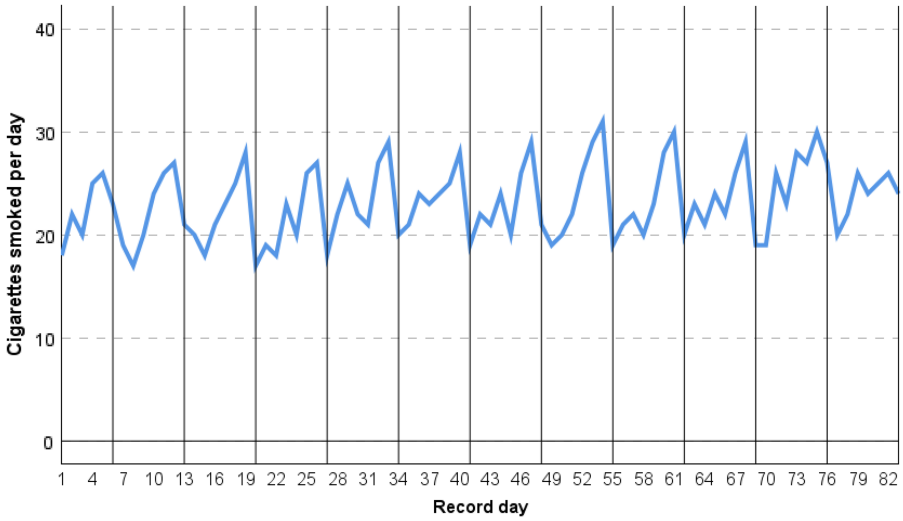
Exploratory Data Analysis. To investigate the AC structure and identify potential seasonal effects, we examine the ACF and PACF up to 21 lags (21 days, or 3 weeks), hypothesizing a 7-day cyclic pattern. Figure 5a and 5b, produced by running Script 3 (R Core Team, 2023b; Wickham, 2016), depict the ACF and PACF for the data set, respectively.

A precision, which is commonly referred to as a "module" in most statistical analysis programs, is called a "library" in R. In this research, we will use both terms interchangeably. The R library for calculating and plotting the ACF and PACF is `stats`, and the basic syntax is as follows:

```
acf(x, lag.max = n, plot = TRUE, or FALSE)
pacf(x, lag.max = n, plot = TRUE, or FALSE)
```

In both functions, `x` is the variable for calculating the ACF and PACF; in our example, it is `data$P48`. The instruction `lag.max` specifies the maximum number of lags to be calculated; the user can

Figure 3
Daily Number of Cigarettes Smoked



Note. The vertical lines represent each Sunday record. See the online article for the color version of this figure.

specify a number `n` of lags, which is 21 in our example. The parameter `plot` can be set to `TRUE` or `FALSE` to enable or disable plotting. Thus, the syntax in our example is:

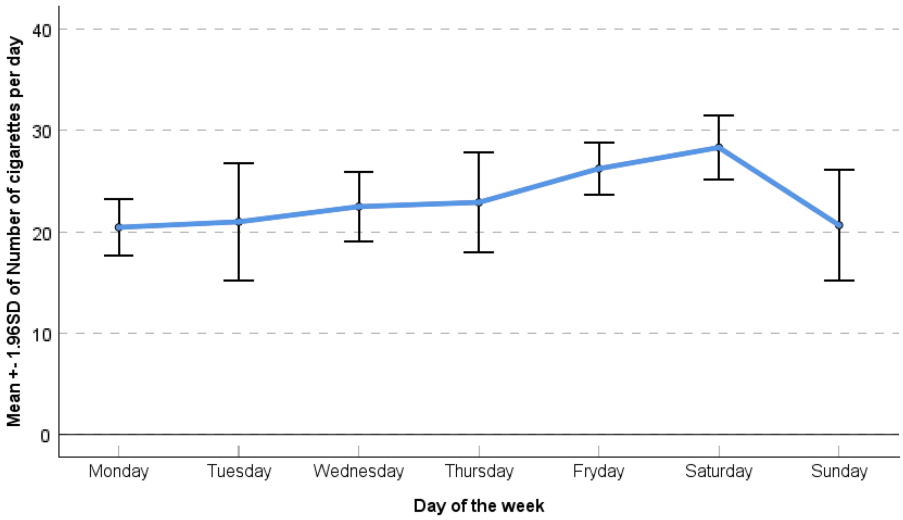
```
acf(data$P48, lag.max = 21, plot = TRUE)
pacf(data$P48, lag.max = 21, plot = TRUE)
```

Note that we occasionally use more complex syntax in the R scripts to obtain nonstandard statistics or more detailed figures.

Preliminary Model Evaluation. The Ljung–Box (L&B) test, applied to 21 lags, yields a χ^2 value of 172.846 with 21 *df*, indicating a *p* value of less than .001. This result suggests that the TS exhibits serial dependence. Furthermore, the Augmented Dickey–Fuller (ADF) test, using 7 lags, shows a statistic of −9.026 with a *p* value of less than .01. This provides strong evidence against the presence of a unit root, suggesting that the series is likely stationary.

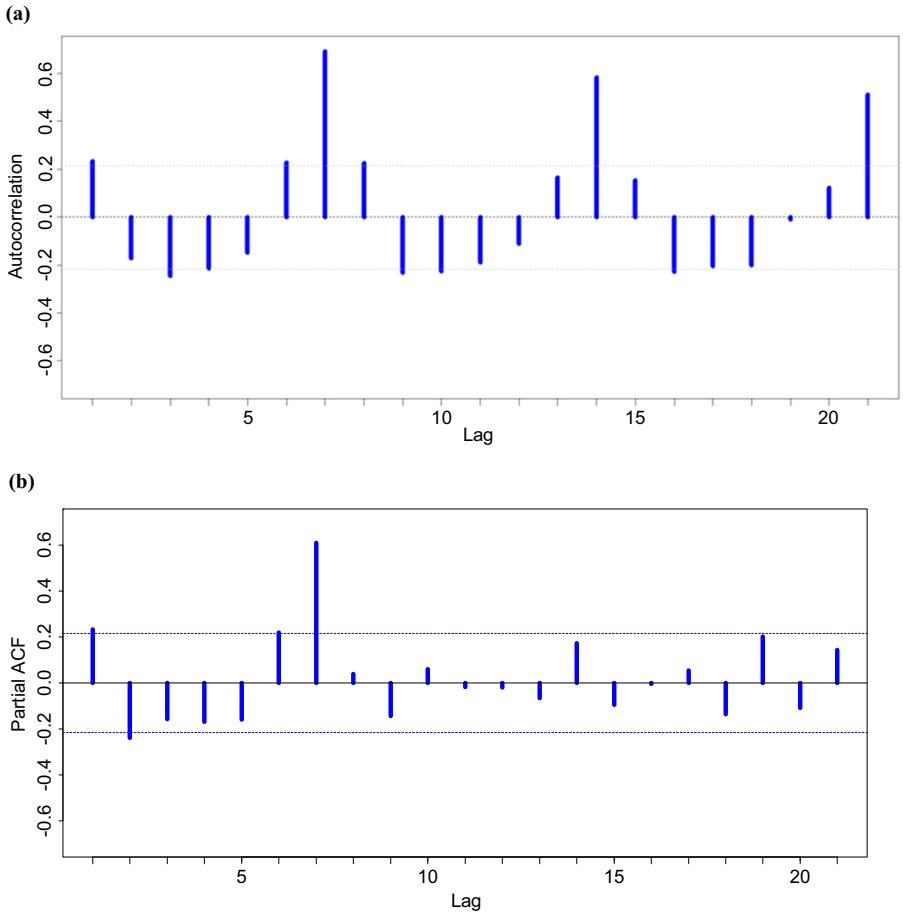
The ACF and PACF analyses indicate significant correlations at lags 1 and 7. Although lag 14 is not significant in the PACF, it is included based on the seasonal hypothesis, leading to the proposed AR(1, 7, 14) model, or $AR(1)(2)_7$. This approach, which incorporates researcher-driven hypotheses while allowing for additional parameters when uncertainty exists, helps prevent the omission of potentially significant IV influences on the DV, Y_t (Anderson, 2011).

Figure 4
Mean ± 1.96 SD of Cigarettes Smoked for Each Day of the Week



Note. Vertical error bars represent the mean ± 1.96 *SD* of the number of cigarettes smoked per day of the week. See the online article for the color version of this figure.

Figure 5
ACF and PACF of the Original P48 Data



Note. Panel (a) shows the Autocorrelation function of P48; Panel (b) shows the Patial ACF of P48. ACF = autocorrelation function; PACF = partial ACF. See the online article for the color version of this figure.

While Box and Jenkins models typically include lags where both ACF and PACF are significant (thus, a Box and Jenkins analyst would have included, according to Figure 5, lags 1 and 7; however, in parameter estimation, we found that lags 7 and 14 are significant, while lag 1 is not), we recommend incorporating lags if at least one function shows significance, provided they align with the proposed hypotheses.

This procedure, known as "overparameterization," should be hypothesis-driven, avoiding the inclusion of unsupported parameters (Campos et al., 2005; Hendry, 2024). Nonsignificant parameters can always be removed in a subsequent estimation phase. The complete syntax for the L&B and ADF tests is available in R Script 4.

The basic syntax for L&B test:

```
Box.test(x, lag = n, type = ("Box-Pierce", or "Ljung-Box")
```

If we have in account the previous attributions for the ACF and PACF syntax, our basic syntax would be:

```
Box.test(data$P48, lag = 21, type = ("Ljung-Box")
```

To obtain the ADF, we need to download the `tseries` library and use the following syntax:

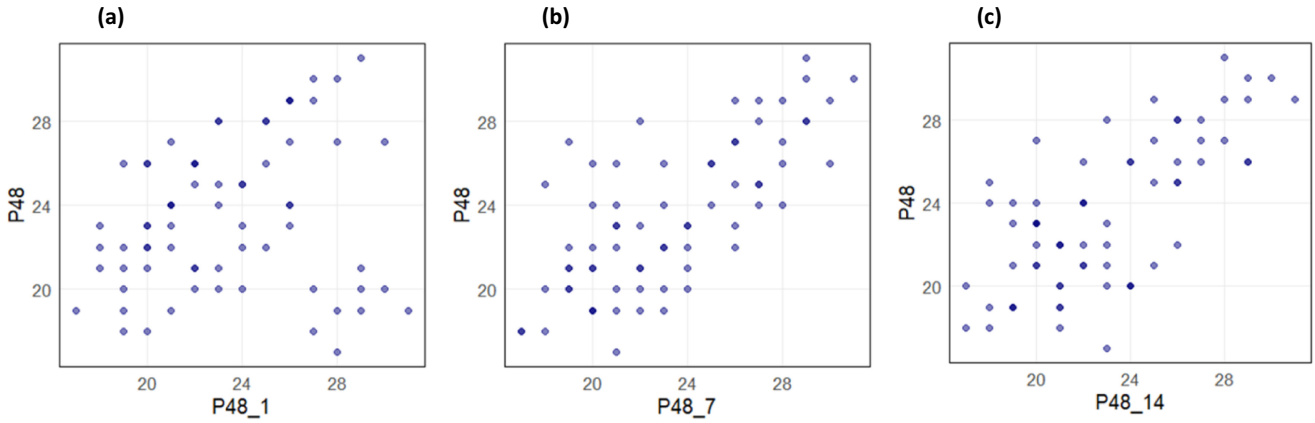
```
adf_test_results <- adf.test(data$P48)
```

A frequently overlooked aspect of AR TS models is their linearity, which implies a straight relationship between the IVs and DV. One way to visualize this relationship is through a scatterplot, or bivariate point cloud, showing the DV against the potential IVs. In Figure 6, we present scatterplots between Y_t (DV), and Y_{t-1}, Y_{t-7} and Y_{t-14} (IVs). The ACF showed that all three correlations are statistically significant, and we confirm in Figure 6 that these relations are linear. However, in Figure 6a, we observe that the relationship between Y_t and Y_{t-1} is the least linear of the three. R Script 5 provides the syntax for generating Figure 6.

Parameters Estimation

To estimate the AR model with lags of 1, 7, and 14 days, or ARIMA(1,0,0)(2,0,0)$_7$ model, we utilized three different procedures: (a) The default ARIMA system, or ARIMA(1,0,0)(2,0,0)$_7$ model, in R Script 6a, which uses ML estimation. (b) OLS for

Figure 6
Scatterplots of Y_t With Lagged Values



Note. Panel (a) shows the P48 versus P48_1, or Y_t versus Y_{t-1}; Panel (b) shows the P48 versus P48_7, or Y_t versus $Y_{t_1}-7$; and Panel (c) shows the P48 versus P48_14, or Y_t versus Y_{t-14}. See the online article for the color version of this figure.

parameter estimation, as documented in R Script 6b, using "temporary" lagged variables (not recorded in a data file). This method was chosen for two reasons: firstly, to allow comparison between ML and OLS, and secondly, because OLS is the BLUE under conditions of normality, homoskedasticity, and no AC among residuals (Kitagawa, 2021). (c) We also used transformed and saved lagged variables in R Script 6c, but with ML estimation. The three methods yielded the same results. Therefore, the model's fit to the smoking data from participant P48 is reported in Table 1, using the "forecast" library in R (Trapletti et al., 2024).

ARIMA Estimation. The basic syntax in R Script 6a is as follows:

```
model <- Arima(data, order = c(p,d,q),
 seasonal = list(order = c(P,D,Q), period = S),
```

in this syntax, the "`order =`" terms correspond to the general ARIMA (p, d, q) (P, D, Q) S model, our specific syntax is:

```
arima_model <- Arima(ts_data, order = c(1, 0, 0),
  seasonal = list(order = c(2, 0, 0), period = 7)
```

OLS Estimation. In R Script 6b, using OLS estimation and "temporary" lagged variables, the basic syntax for creating a lag 7 temporary variable is:

```
lag7 <- c(rep(NA, 7), head(ts_data, -7))
```

and for the OLS regression, the syntax is:

```
model <- lm(ts_data ~ lag1 + lag7 + lag14)
```

Table 1
Ordinary Least Squares Estimation of AR(1, 7, 14) Model in R

Parameter	Estimate	*SE*	*t*	Pr(>|*t*|)
Intercept	4.368	2.472	1.767	0.082
Y_{t-1}	0.035	0.082	0.429	0.669
Y_{t-7}	0.430	0.118	3.647	0.001
Y_{t-14}	0.361	0.122	2.944	0.004

Note. Dependent variable: Y_t, daily smoking of P48. This model corresponds to Equation 5 in the main text. AR = autoregressive.

Estimation With Generated Lagged Variables. In R Script 6c, with ML estimation and generated lagged variables, it is necessary to load the `lmtest` library; the regression syntax is:

```
xreg <- cbind(lag1, lag7, lag14)
```

Observe that in R Script 6a, the syntax represents the ARIMA(1, 0, 0)(2, 0, 0)$_7$ model, while in Scripts 5b and 5c, the instructions pertain to the AR(1, 7, 14) model. Despite these different expressions, the two models are equivalent and yield identical results under the same estimation conditions.

Results

The results of R Scripts 6a, 6b, and 6c match and are shown in Table 1, and are also identical to those of R Script 10.

The model presented in Table 1 is as follows:

$$Y_t = 4.368 + .035Y_{t-1} + .430Y_{t-7} + .361Y_{t-14} + e_t. \quad (5)$$

The model's overall fit, $F(3, 65) = 29.78$, $p < .001$, $R^2 = .579$, $\text{Var}(e_t) = 5.244$, Akaike information criterion (AIC) = 121.353, indicates a substantial explanation of the variance in smoking data. Notably, the lag at Y_{t-1} did not show statistical significance ($p = .669$), contrary to expectations based on its ACF and PACF significance in Figure 5a and 5b. However, this lag was retained in the model as a hypothesis. Lags at Y_{t-7} and Y_{t-14} were found to be significant in the regression equation ($p < .001$ and $p = .004$, respectively), despite the PACF for Y_{t-14} not showing significance in Figure 5b. These findings underscore the utility of including additional parameters in the estimation process when in doubt, a practice referred to as overparameterizing, as ACF and PACF results are primarily indicative (Basu, 2020; Freedman, 2012a; Greene, 2012). The advantage of overparameterization is that, on the one hand, it provides the opportunity for lags with questionable significance in exploratory analysis through ACF and PACF to potentially resolve their significance in the final model. On the other hand, it allows for IVs significant in ACF and PACF with high correlation with other IVs to become nonsignificant when estimating the model

parameters. The primary error in TS analysis, and in any regression equation, is the omission of any relevant variable. Not including relevant variables, while adding irrelevant variables to the model, will result in biased and inconsistent estimates of the latter. This bias persists regardless of the sample size (Basu, 2020; Kmenta, 1971).

If we remove the variable Y_{t-1}, the results would be very similar to those in Table 1. The equation without Y_{t-1} would now be:

$$Y_t = 5.006 + .433Y_{t-7} + .364Y_{t-14} + e_t, \quad (6)$$

with $F(2,\ 66) = 45.13,\ p < .001,\ R^2 = .578,\ \text{Var}(e_t) = 5.262$, AIC = 119.148. The coefficient of Y_{t-7} (.433) has a $p < .001$, while the coefficient of Y_{t-14} has a p value of .004. This indicates that when a nonsignificant variable is removed, the significant variables barely change (there are no biases in the estimation of the parameter coefficients or their respective probabilities), and the explanatory power of the model remains the same (as indicated by the probability of the F test and the R^2).

However, if we followed the rule that a lag must be included when its ACF and PACF are significant, as shown in Figure 5a and 5b, lags 1 and 7 are significant in both functions. Thus, including Y_{t-1} and Y_{t-7} in the explicative model, we get:

$$Y_t = 5.600 + .066Y_{t-1} + .700Y_{t-7} + e_t, \quad (7)$$

being $F(2,\ 73) = 42.12,\ p < .001,\ R^2 = .536,\ \text{Var}(e_t) = 5.799$, AIC = 138.546. This means that the explanatory capability of Equation 7 is smaller than that of Equations 5 or 6, as indicated primarily by its R^2. The Y_{t-1} coefficient (.066) has $p = .422$, not significant, and the Y_{t-7} coefficient (.700), $p < .001$. Additionally, if we compare Equation 7 with the equations where the significant variables Y_{t-7} and Y_{t-14} are included, the coefficient of .700 is much larger (biased?) than in Equations 5 or 6.

In R Script 6b, the `'na.action = na.exclude'` syntax was implemented to ensure no data were omitted from the estimation of forecasted values and residuals. This step is crucial because, by default, R would eliminate any row containing missing values, which, without this specification, would result in the loss of the first 14 original rows of data in the new file with the estimated values, saved as "P48_TS_OLS_Frkst&Rsids_AR_1_7_14.rds."

Diagnostic Assessment. By analyzing the ACF and PACF of residuals from the "P48_TS_OLS_Frkst&Rsids_AR_1_7_14.rds" data set, which contains forecasted and residual values based on the model depicted in Table 1, we observe diagnostic results. According to the procedure outlined in R Script 7, the outcomes are displayed in Figure 7a and 7b.

The L&B test for 21 lags yields $\chi^2(21) = 18.147,\ p = .640$, indicating no serial dependency within the residual TS. The Shapiro–Wilk (S&W) test for normality returns a statistic of .986 with a p value of .660, affirming the residuals' normal distribution. To assess homoscedasticity, we employ the studentized Breusch–Pagan (B&P) test, examining the variance equality of residuals across the different days of the week. The test statistic stands at $\chi^2(6) = 7.379$, yielding a p value of .287. This supports the presumption of variance homogeneity among the residuals.

Given these diagnostic checks, with the absence of AC, adherence to normality, and uniform variance, we conclude that the residuals are "white noise." The processes for conducting the L&B, S&W, and B&P tests are detailed in R Script 8.

Model Refinement and Validation. Notably, in Figure 7a, the ACF at lag 19 emerges as significant, potentially attributable to random variation. Given the PACF at lag 19 significance, and the L&B test for lag 21 also returns a nonsignificant outcome, we might disregard this anomaly. Had both the ACF and PACF at lag 19 shown significance, it might prompt reconsideration for including an AR(19) term in the model. However, our foundational hypothesis posits that an individual's smoking behavior is influenced by recent and weekly seasonal lagged (7, 14, 21 … days) smoking patterns, suggesting the significance at lag 19 in the residuals may be incidental, thus maintaining the AR(1, 7, 14) model as originally formulated.

A different situation would arise if significance appeared at a lag immediately adjacent to those specified in our hypothesis, for instance, at lag 2. If both the ACF and PACF at lag 2 were significant, it would be advisable to include this lag as an additional predictor in the model. In that case, the revised specification would become $Y_t = f(Y_{t-1},\ Y_{t-2},\ Y_{t-7},\ Y_{t-14})$, while imposing a constraint such that $b_1 \geqslant b_2$, in line with the assumption that more recent observations exert greater influence on current behavior.

Model validation can be approached through various methods, including invertibility and unit root tests. Yet, a straightforward validation technique for TS involves examining the LTT (Engle & Granger, 1987). The LTT test has been primarily developed within the context of multivariate TS analysis. Therefore, one might question what constitutes a valid LTT value for a univariate TS. A value within the confidence interval of the mean plus or minus 1.96 *SD*s, $\overline{Y_t} \pm 1.96 \times SD(Y_t)$, would be considered acceptable for the LTT of a univariate series. In this research, the LTT result, detailed in Appendix B, is 25.103, an acceptable figure, thereby endorsing the statistical model presented in Table 1 and Equation B1 in Appendix B.

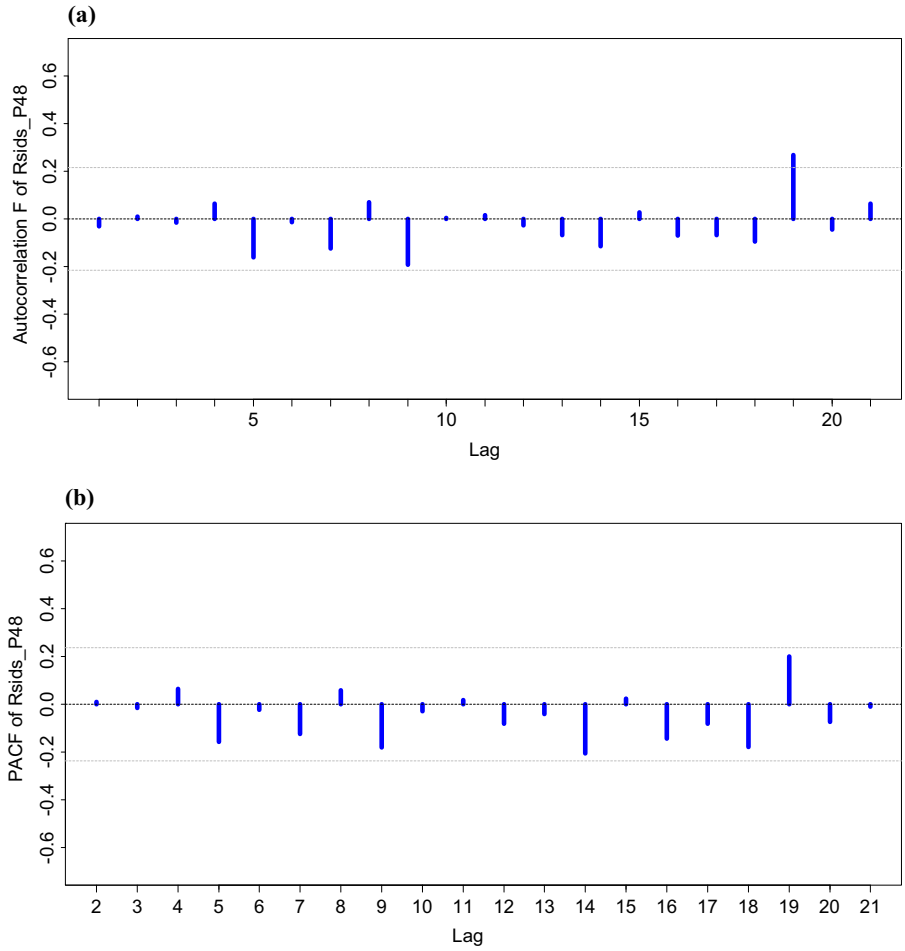
Although the AR(1) component of our model is nonsignificant, it remains part of the model to support our hypotheses. Thus, the AR(1, 7, 14) configuration effectively takes into account behavioral characteristics of the participant's daily smoking, where the number of cigarettes smoked on a given day (Y_t) is influenced by the prior day's consumption (Y_{t-1}) as well as by smoking patterns 1 and 2 weeks prior (Y_{t-7} and Y_{t-14}), reinforcing the hypothesis of recurrent behavior based on immediate and seasonal influences.

In conclusion, the model outlined in Table 1 is deemed appropriate for describing the smoking behavior of the participant denoted as P48.

Interpretation of Coefficients. TS models are a special case of regression. In the regression equation $Y = b_1X_1 + b_2X_2 + e$, the coefficients b_1 and b_2 represent the partial effects of X_1 and X_2 on Y, while controlling for the other variable. Specifically, b_1 indicates the expected change in Y for a one-unit increase in X_1, holding X_2 constant, while b_2 represents the expected change in Y for a one-unit increase in X_2, holding X_1 constant. For example, in a study examining the effects of study hours (X_1) and hours of transport from home to university (X_2) on exam scores (Y), if $b_1 = 2$, it means that for every additional hour of study, the exam score increases by 2 points, assuming transportation time remains constant. Conversely, if $b_2 = -1.5$, it implies that for every additional hour spent commuting (X_2), the exam score decreases by 1.5 points, assuming study hours ($X1$) remain constant. Thus, b_1 and b_2 reflect the unique contribution of each predictor to Y, accounting for the presence of the other

Figure 7

ACF and PACF of Residuals From the AR(1,7,14) Model, Equation 5



Note. Panel (a) shows the ACF of P48 residuals; Panel (b) shows the PACF of P48 residuals. ACF = autocorrelation function; PACF = partial ACF; AR = autoregressive. See the online article for the color version of this figure.

variable. The same principles apply to the regression coefficients in a TS model.

If Equation 2 ($Y_t = 4.368 + .035Y_{t-1} + .430Y_{t-7} + .361Y_{t-14} + e_t$) is considered valid, the interpretation of its coefficients is as follows: (a) The coefficient of Y_{t-1} (.035) suggests that, holding all other factors constant, a one-unit increase in Y at time $t-1$ (the previous day) is associated with a .035-unit increase in Y at time t. This indicates a weak positive short-term dependency, implying that the amount this person smokes on a given day slightly influences their smoking behavior the following day. This effect persists throughout the series (from $t = 1$ to $t = 83$). However, it is important to note that this coefficient is not statistically significant, meaning we cannot confidently assert that this effect is different from zero. (b) Regarding the interpretation of the coefficient for Y_{t-7} (.430), it indicates that, holding all other factors constant, a one-unit increase in Y at time $t-7$ (1 week prior) is associated with a .430-unit increase in Y at time t. In the context of smoking behavior, this suggests a significant weekly pattern: if our individual smokes an additional cigarette on a given day, it is likely that he will smoke approximately .430 more cigarettes 7 days later. This finding highlights a strong weekly seasonality, implying that smoking habits tend to repeat on a weekly basis. For instance, if our person smokes 10 cigarettes (1×10) on a Tuesday, it is expected that the following Tuesday, he will smoke, on average, 4.3 additional cigarettes ($.430 \times 10$). This pattern is consistent across all days of the week (Mondays, Tuesdays, …, or Sundays) throughout the entire series. (c) The coefficient of Y_{t-14} (.361) indicates that, holding all other factors constant, a one-unit increase in Y at time $t-14$ (the same day, 2 weeks prior) is associated with a .361-unit increase in Y at time t. This suggests a biweekly (14 days) cyclical pattern, where the value of Y is significantly influenced by its value from 2 weeks earlier. In practical terms, for the person of our research, this means that the number of cigarettes smoked on a given day (e.g., a Thursday) affects the number of cigarettes smoked two Thursdays later.

Overall, these findings indicate that both short-term (Y_{t-1}) and seasonal (Y_{t-7}, Y_{t-14}) dependencies are instrumental in forecasting Y_t. The presence of weekly and biweekly lags suggests a recurring

temporal structure within the data, which is essential for comprehending and predicting patterns in the TS. While TS models are typically expressed functionally, as in our case, $Y_t = f(Y_{t-1}, Y_{t-7}, Y_{t-14})$, they can also be interpreted causally. That is, an individual's smoking behavior on a given day (Y_t) directly influences their consumption the following day (Y_{t+1}), 7 days later (Y_{t+7}), and 14 days later (Y_{t+14}).

Regression Estimation With `R dplyr` *Library*

This section, alongside its SPSS counterpart, forms the crux of our article, demanding clear comprehension. In the preceding segment on TS estimation, we utilized the "`forecast`" package in R for automated system estimation. In previous analyses, the values of Y_t were automatically lagged, so the analyst did not observe the variable lagging process. However, in this section, variables will be manually lagged, added, and saved in the original data file. Yet, considering TS analysis as a specialized form of regression, where the DV, Y_t, is influenced by its past values (Y_{t-1}, Y_{t-2}, Y_{t-3}, …), transforms Y_t into both the DV and, when properly lagged, its own IV. This realization implies that with the ability to lag Y_t values, specialized TS software becomes unnecessary; instead, we can employ ordinary regression analysis with any basic linear regression tool available in standard statistical packages.

The initial steps for analyzing the smoking data of participant P48 mirror those outlined earlier: data exploration (Figures 3 and 4), identifying characteristics via ACF and PACF of the original data (Figure 5), conducting the L&B test to assess serial dependency, and applying the ADF test for stationarity checks. Opting for the AR(1, 7, 14) model estimation, our approach simplifies the process. We begin by lagging P48's smoking data by 1, 7, and 14 days, preparing for regression analysis without necessitating specialized TS estimation tools.

Lagging Variables With R. To prepare our data for regression analysis, we lag the P48 variable from the "P48.rds" file, creating lags of 1, 7, and 14 days, named "P48_1," "P48_7," and "P48_14," respectively. These modified variables are saved in a new file, "P48_lagd.rds," using R Script 9. A crucial line in this script ensures no missing values are removed during this process (`data <- data %>%`), as R typically deletes entire rows containing any missing values by default.

The basic syntax for lagging variables in R, using the variable P48 to generate P48 lagged by 1 day (name: `P48_1`), 7 days (`P48_7`), and 14 days (`P48_14`), is as follows:

```
data <- data %>% # Do not delete line with missing
data mutate(P48_1 = lag(P48, n = 1), P48_7 =
lag(P48, n = 7), P48_14 = lag(P48, n = 14)) # Lagged
variables: 1, 7 and 14
```

Note that in R programming, the symbol "#" denotes a comment inside that line. Any text in a line following this symbol is considered a notation and is not processed as part of the R syntax.

Distinguishing Between Original and Lagged Data. The original "P48.rds" and the newly lagged "P48_lagd.rds" files differ primarily in the transformation applied to create temporal lags for regression analysis, used in R Script 6. Figure 8 visually depicts this transformation. For example, in Figure 8a, P48 on Day 1 (ID_Day = 1, or $t = 1$) is highlighted in yellow, showing a value of 18 ($Y_1 = 18$). Figure 8b shows how this value shifts to the second line in the "P48_1" column (representing Y_{t-1}), indicating lagging by 1 day. This lagging process is illustrated by analyzing the functional relationship $Y_t = f(Y_{t-1}, Y_{t-7}, Y_{t-14})$, using the corresponding lagged transformations of Y_t performed in Script 6c. For instance, the value at Line 19 ($Y_{19} = 28$) is influenced by the values from $Y_{18} = 25$, $Y_{12} = 27$, and $Y_5 = 26$.

Handling Missing Values and Data Shifts. An important observation from Figure 8b is the maintained row count (83), identical to Figure 8a. Lagging shifts the data set backward, introducing missing values at the series start without affecting the series' end, contrary to some software behaviors that omit rows with missing data, which can lead to errors when forecasting data through "handmade" procedures.

This "cutting" of values significantly impacts parameter estimation, especially in multivariate analysis, where AR univariate models are a specialized form. Utilizing "listwise deletion" for AR(1, 7, 14) model analysis means that only complete data from Rows 15–83 (68 data points) are considered, excluding initial rows due to missing lagged values.

This paragraph has emphasized the methodological nuances of preparing TS data for regression analysis in R, highlighting the importance of careful data handling and the implications of different analytical choices.

Results: Regression Parameter Estimation With R. After identifying the AR(1, 7, 14) model using R Scripts 1 through 8, we proceed with regression analysis using R Script 10, where $Y_t = f(Y_{t-1}, Y_{t-7}, Y_{t-14})$, or equivalently, $P48 = f(P48_1, P48_7, P48_14)$. The regression results align with those presented in Table 1, confirming our diagnostic checks (ACF and PACF of residuals, L&B test for serial dependence, S&W test for normality, and B&P test for homoscedasticity) yield consistent findings (see Figure 7).

This analysis highlights that understanding the DV P48's lagged values and applying OLS regression suffices for AR model estimation. In contrast, using modules like "`forecast`" or "`ts`" in R package for TS analysis might overwhelm beginners with a "black box" impression, obscuring the underlying statistical processes.

Forecasting. Given their nature as a specialized regression case, it is straightforward forecasting with AR models. Utilizing the coefficients from Table 1 and the data for Day 19 as depicted in Figure 8b or Figure 7c, we articulate the general equation for the AR(1, 7, 14) model as follows:

$$Y_t = 4.368 + .035Y_{t-1} + .430Y_{t-7} + .361Y_{t-14} + e_t. \quad (8)$$

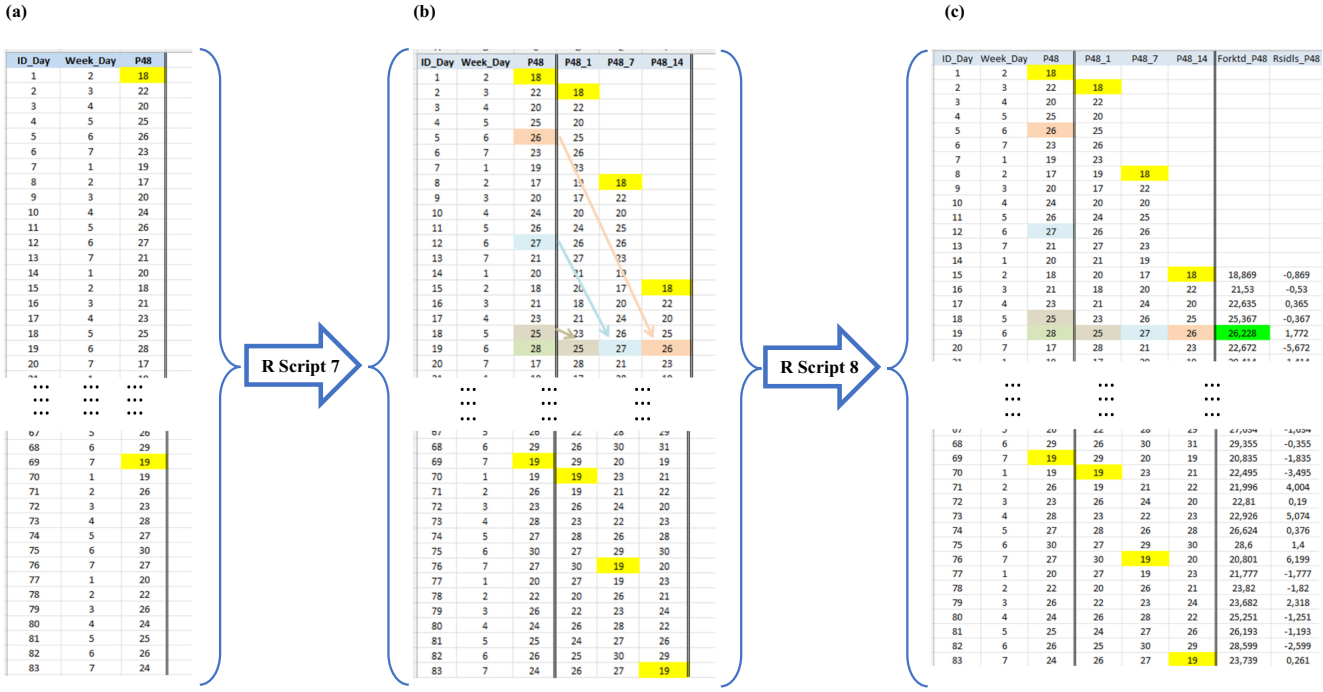
Focusing on Day 19, with $Y_{19} = 28$ (or Y_t), $Y_{18} = 25$ (or Y_{t-1}), $Y_{12} = 27$ (Y_{t-7}), and $Y_5 = 26$ (Y_{t-14}), the forecasted value $\widehat{Y}_{19}$ is calculated as:

$$\widehat{Y}_{19} = 4.368 + .035 \times 25 + .430 \times 27 + .361 \times 26 = 26.228. \quad (9)$$

This results in an error for Day 19, e_{19}, of 1.772, being $Y_{19} = \widehat{Y}_{19} + e_{19}$ (i.e., $e_{19} = 28 - 26.228 = 1.772$), corresponding to the observed values for this day. It is important to note that manually calculating Y_t might introduce minor rounding discrepancies due to the extended decimal places of coefficients in Equation 8 not fully represented here. For the calculation of confidence intervals for forecast values, refer to Hyndman and Athanasopoulos (2018) and Shumway and Stoffer (2017).

Figure 8

Construction of Data Sets for TS Regression Modeling



Note. Panel (a) shows the original 'P48' data; Panel (b) shows the lagged dataset ('P48_lagd'): original + lagged predictors; Panel (c) shows the extended dataset ('P48_lagd_lin_reg'): Includes forecasted values and residuals. TS = time series. See the online article for the color version of this figure.

This approach underlines the efficiency of AR models in forecasting, simplifying the predictive process by leveraging regression coefficients and past data values.

TS With the IBM SPSS Package

IBM SPSS, with a good interface for Windows environments and an easy-to-use syntax editor, offers good usability on both platforms. To streamline comparisons, we will highlight key distinctions between R and SPSS (Tokunaga, 2021). Detailed instructions and syntax for replicating our analysis in the R package are provided in the "SPSS Scripts.docx" document. To ensure consistency, we have aligned script order with that of the "R Scripts.docx," facilitating direct comparisons between the two packages. However, modifications were necessary in SPSS Script 6 and beyond due to estimation discrepancies encountered in the SPSS TSMODEL module. Additionally, procedural and theoretical aspects addressed in "Data Analysis 1: With the R package" will not be repeated in this section to avoid redundancy.

Identification

In SPSS Script 1a, we detail the syntax steps, and we briefly introduce the graphical interface to convert data from Excel format ("*.xlsx") to SPSS format ("*.sav"), and the reverse process is outlined in SPSS Script 1b. Following the successful import into the "P48.sav" file in SPSS, creating Figure 3 is straightforward with the SPSS syntax provided in SPSS Script 2a. Similarly, Figure 4 can be generated using the instructions in SPSS Script 2b. However, both figures require additional adjustments via the SPSS interactive interface to finalize their presentation.

SPSS Script 3 includes syntax for computing the ACF and PACF for the original data set. The basic general SPSS syntax is:

```
ACF VARIABLES= V1 V2 V3 /* V1 V2 V3 … are the name of variables
 /MXAUTO n     /*(1) Maximum number of lags to analyze: n
 /PACF.    /*(2) Displays and plots PACFs
```

/*(1) We recommend setting the '/MXAUTO' parameter to at least three seasons (n= 3*7 = 21 lags or days).

/*(2) Use the PACF subcommand to display and plot sample PACF for each series named on the ACF command (V1 V2 V3).

In our example, we analyze only the variable P48, with 21 lags, being the syntax:

```
ACF VARIABLES=P48
  /MXAUTO 21   /* We recommend at least three seasons (3*7 = 21 lags, or days)
  /PACF.   /* The last line in SPSS syntax must end with a period (".")
```

We have shown that in R, blank lines or information within the same paragraph can be left, as long as it is preceded by the "#" sign, so that R will ignore that text; however, the same cannot be done in SPSS. Therefore, it is very important in SPSS to keep paragraphs separate without line breaks, although annotations can be added at the end of any line, and they must be preceded by a "/*" sign so that the SPSS

program does not consider them as part of its syntax. See examples of scripts with R and SPSS syntax in the corresponding files.

Furthermore, SPSS Script 4 is dedicated to performing the L&B and ADF tests, essential for diagnosing the TS data. These steps are critical for identifying the underlying patterns and statistical properties of the data set, facilitating a comprehensive analysis within the SPSS environment. The L&B test syntax is implicitly included in the "ACF VARIABLES" paragraph. To run the ADF tests, the SPSS "Modeler" module is required.

Using the SPSS Script 5, you can obtain the syntax to check the linearity between the DV: Y_t, and the IVs: Y_{t-1}, $Y_{t-7,}$ and Y_{t-14}. The resulting figures are substantially the same as those in Figure 6 with R.

Estimation Through TS With SPSS `TSMODEL` *Module*

Data Analysis. The SPSS `TSMODEL` module does not support OLS estimation; instead, we utilized the default ML estimation. To set the ARIMA(1,0,0)(2,0,0)$_7$model, you first need to run the Windows interface instructions in Script 6b, Steps 1 and 2, to determine the seasonality of 7 days. After this, you can generate the syntax with the following basic instructions:

```
TSMODEL
 /MODEL DEPENDENT=V1 V2 /* DVs: V1 V2…
 /ARIMA AR=[p] DIFF=d MA=[q] ARSEASONAL=[P]
DIFFSEASONAL=D MASEASONAL=[Q] /* ARIMA(p,d,q)(P,
D,Q), not S
 CONSTANT=YES. /* Include equation constant
```

Note that in the previous paragraph, there is no information about the seasonality, S, of the series. However, SPSS has in memory that in the second step, we specified that the "Periodicity at higher level" is 7 (or 7 days of seasonality). The essential syntax for running our variable is:

```
TSMODEL
 /MODEL DEPENDENT=P48
 /ARIMA AR=[1] DIFF=0 MA=[0] ARSEASONAL=[1,2]
DIFFSEASONAL=0 MASEASONAL=[0]
 CONSTANT=YES.
```

It is very important to write `ARSEASONAL=[1,2]` in the previous syntax to obtain parameters corresponding to Y_{t-7} and Y_{t-14}. If you write `ARSEASONAL=[2]`, you will only get the parameter corresponding to Y_{t-14}. The results of the complete SPSS Script 6a are presented in Table 2, where we have highlighted in grey color the estimated value of the constant.

Table 2
Results of SARIMA(1, 0, 0)(2, 0, 0)$_7$ Time Series Model in `TSMODEL` *SPSS*

Parameter	Estimate	*SE*	*t*	Pr(>|*t*|)
Intercept	23.179[a]	0.960	24.136	<0.001
Y_{t-1}	0.061	0.113	0.543	0.588
Y_{t-7}	0.465	0.111	4.190	<0.001
Y_{t-14}	0.329	0.119	2.763	0.007

Note. Dependent variable: Y_t, daily smoking rate of P48. This model corresponds to Equation C1 in Appendix C. SARIMA = seasonal, autoregressive, integrated, and moving average.
[a] We have highlighted the erroneous estimated value of the intercept.

The basic general instructions for the AR(1, 7, 14) model is in the SPSS Script 6c, being the basic syntax:

```
TSMODEL
 /MODEL DEPENDENT= V1 V2 /* Define and fit the DVs (V1
V2…)
 /ARIMA AR=[p] DIFF=d MA=[q] ARSEASONAL=[P]
DIFFSEASONAL=D MASEASONAL=[Q] /* ARIMA(1,0,0)
(2,0,0) model.
```

And the indispensable syntax for the variable P48, AR(1, 7, 14) model, is:

```
TSMODEL
 /MODEL DEPENDENT=P48
 /ARIMA AR=[1, 7, 14] DIFF=0 MA=[0]
ARSEASONAL=[0] DIFFSEASONAL=0 MASEASONAL=[0]
/* AR(1, 7, 14)
```

Note that the AR instruction for lags 1, 7, and 14 of P48 is `AR=[1, 7, 14]`. The results are the same as those with the SPSS Script 6b, as shown in Table 2.

Upon comparing Table 2 with Table 1, minor discrepancies in the coefficients for the IVs: Y_{t-1}, $Y_{t-7,}$ and Y_{t-14} are observed, alongside a significant difference in the intercept values: 4.368 in Table 1 versus 23.179 in Table 2. We have highlighted with grey color the estimated value of the intercept. We contend that the results depicted in Table 2, derived from SPSS, are inaccurate. A concise critique of the errors associated with the `TSMODEL` SPSS estimation is available in Appendix C.

Lagging Variables in SPSS. With the aim of lagging variable "P48," generating variables "P48_1," "P48_7," and "P48_14" (representing Y_{t-1}, Y_{t-7}, and Y_{t-14}, respectively), following the steps outlined in SPSS Script 6, you will find "P48_lagd.sav" file, which displays a configuration similar to that shown in Figure 8b. The basic syntax for our data is:

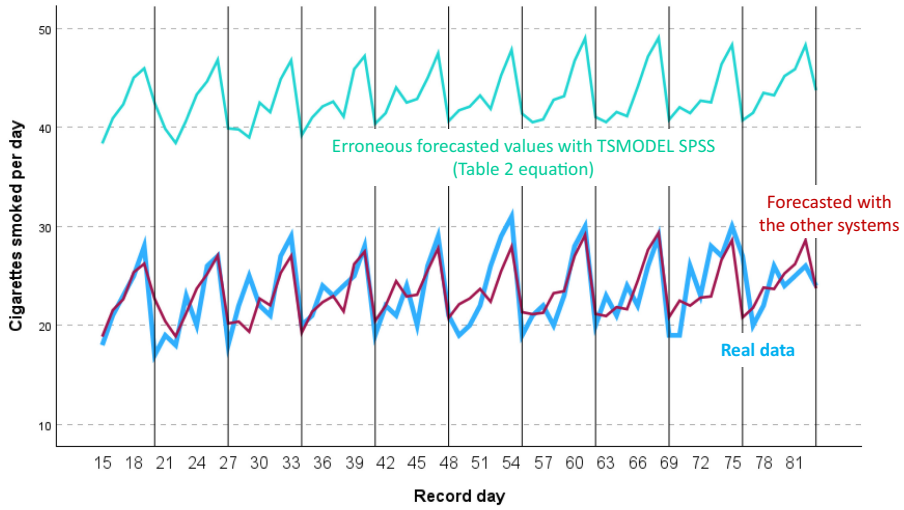
```
CREATE
 /P48_1=LAG(P48 1) /*(1)
 /P48_7=LAG(P48 7)
 /P48_14=LAG(P48 14).
/*(1) New variable name: (P48_1) = Lag((original vari-
able, P48), number of lags: 1))
```

This arrangement ensures the data set is ready for analysis using the SPSS `COMPUTE` or `REGRESSION` modules.

Calculating Forecasted Values Agree Table 2 Equation. Furthermore, forecasted and residual values, based on the equation from Table 2, as it is exposed on SPSS Script 7, have been compiled in the "P48_TS_Tbl2.sav" file. Figure 9 contrasts the original data set against both the flawed forecasted values from the `TSMODEL` SPSS module and accurate forecasts obtained via alternative methods, including TS and OLS regression in R, as well as OLS regression in SPSS. Due to the noted estimation inaccuracies, data processed through the `TSMODEL` SPSS module are explained in Equation C3 in Appendix C. In this appendix, we manually calculate the expected values of Y_t from Table 2, revealing a substantial discrepancy in the intercept estimate. We isolate there the intercept b_0 and finally, manually estimate the predicted values of the DV Y_t, checking if the mean of the errors is significantly different from zero.

Figure 9

Observed and Forecasted Daily Cigarette Consumption Using TSMODEL (SPSS) and Alternative Estimation Systems



Note. Vertical lines mark Sundays. See the online article for the color version of this figure.

Regression Estimation With SPSS `REGRESSION` *Module*

To conduct a temporary regression analysis using SPSS, it is required to lag the P48 variable by one, seven, and 14 units, as outlined in SPSS Script 7.

Regression Data Analysis With SPSS. To conduct regression analysis in SPSS, load the "P48_lagd.sav" file and execute the syntax provided in SPSS Script 9a, the basic syntax:

```
REGRESSION
    /STATISTICS COEFF OUTS R ANOVA /* overall fit
indexes
   /DEPENDENT P48 /* DV: P48
  /METHOD=ENTER P48_1 P48_7 P48_14 /* IVs: P48_1
P48_7 P48_14
    /SAVE PRED RESID. /* printing forecasted and
residual values inside the file
```

This process yields results consistent with those displayed in Table 1. Notably, at the conclusion of the analysis within the "P48_lagd.sav" file, two additional variables emerge: "PRE_1" (the predicted values) and "RES_1" (the residuals), which align with the values depicted in Figure 8c. SPSS Script 9b guides you through renaming these variables and saving them alongside the original data in a newly created file, "P48_lagd_lin_reg.sav."

Diagnostic Assessment. Diagnostic procedures for the regression model in SPSS are outlined in SPSS Script 10 and SPSS Script 11, focusing on analyzing the residuals ("Rsids_P48") from the "P48_lagd_lin_reg.sav" file. Given that the overall results, forecasted values, and residuals match those obtained using R, the residual diagnostics should also align. Consequently, we perform ACF and PACF analyses on the "Rsids_P48" variable, alongside the L&B test, S&W test for normality, and B&P test for homoscedasticity in SPSS. Due to the absence of the "Modeler" SPSS module on our system, the ADF test is not conducted in our SPSS analysis; however, the outcomes would presumably mirror those from the R analysis.

Data Analysis 3: Straight Line Regression

Method

Since some studies use time or squared time as an IV (Fried et al., 2022; Kuba & Scheibe, 2017; Powers & Young, 2008), to examine how the passage of time affects tobacco consumption, we will conduct an analysis to assess how a straight line (or order one polynomial) model fails to meet two critical properties required for a temporal model: (a) the absence of ACR, which can lead to Type I errors, and (b) the preservation of LTT dynamics.

Based on the above considerations and to compare the results with those obtained in the AR(1,7,14) model, we propose here the hypothesis that the number of cigarettes smoked on any given day is a function of the days elapsed since the beginning of the study ($t = 1, 2, \ldots, 83$), or $Y_t = f(t)$. Furthermore, in line with the hypothesis regarding "reactivity" to observed behavior (McCarthy et al., 2015), it is suggested that participants initially reduced their consumption at the start of the recording period, followed by a gradual increase. Therefore, it is expected that the slope of $Y_t = f(t)$, or b_1 in $Y_t = b_0 + b_1 \times t + e_t$, will be positive (upward trend).

Results

This text will not detail how to perform regression analyses with R or SPSS, as these procedures have been described previously; the syntax for these analyses can be found in R Script 11 and SPSS Script 12. According to the results presented in Table 3, the forecasting equation is: $Y_t = 21.487 + .041 \times t + e_t$, with both coefficients being significant (b_0, $p < .001$ and b_1, $p = .010$). The overall model fit, $F(1, 81) = 6.872$, $p = .010$, $R^2 = .078$, $\text{var}(e_t) = 11.296$, AIC = 204.228, indicates a significant explanation of the variance in smoking data.

Table 3
Results of the Regression Model $Y_t = f(t)$, With OLS Estimation

Parameter	Estimate	*SE*	*t*	Pr(>|*t*|)
Intercept	21.487	0.749	28.682	<0.001
ID_Day	0.041	0.015	2.621	0.010

Note. Dependent variable: Y_t, daily smoking of P48. OLS = ordinary least squares.

Interpretation of Coefficients

The interpretation of these results is straightforward: the participant starts with a theoretical baseline of 21.487 cigarettes smoked on the first day. For each additional day, their smoking increases by 0.041 cigarettes. After 100 days, it is expected that the participant will smoke 25.587 cigarettes ($Y_{100} = 21.487 + .041 \times 100$).

Diagnostic Assessment

To evaluate the adequacy of the model, we will perform the ACF and PACF tests on the residuals of the regression model from Table 3, which corresponds to a linear time trend-only specification ($Y_t = b_0 + b_1 \times t + e_t$), without AR terms. As well as the L&B and the LTT tests. Regarding the ACF and PACF of the residuals, they have a pattern very similar to the original data from P48, as is observed in Figure 5a and 5b. The L&B test for 21 lags yields $\chi^2(21) = 197.458$, $p < .001$, indicating that the residuals are not white noise. Consequently, the model in Table 3 fails to eliminate the serial dependence in the original data, meaning it is not a good explanatory model and likely involves a Type I error. Furthermore, when calculating the LTT of the model in Table 3 using the LTT, it is found that $Y^* = 21.487 + .041 \times \infty = \infty$. This result renders the model unviable (Engle & Granger, 1987; Hendry et al., 1984).

Model Comparison: AR(1, 7, 14) "Versus" Straight Line

Various procedures exist for comparing statistical models to determine the optimal explanatory model for a given DV. While these methods generally produce similar results, we opted for the confounding variables procedure due to its simplicity and intuitive interpretation. Originally developed in epidemiology (Maldonado & Greenland, 1993; Westreich, 2020), the confounding test has established connections to causal inference within statistical theory (Pearl, 2009; Wysocki et al., 2022). This approach relies on the stability of regression coefficients as an indicator of model fit.

There are different versions of the confounding variables test, but when there are two sets of competing IVs in regression, such as AR versus straight line (order one polynomial), it is necessary to run each set of "crude" IVs separately, obtaining a crude regression model for each. Finally, both sets of IVs are run together in the "adjusted" test. The IVs that are similar in both the adjusted and crude regression models are the true IVs explaining the DV. If the crude and adjusted IV estimators differ significantly, they are confounding IVs. This method does not require a formal statistical test but rather a comparison of estimators.

The confounding variables test can be implemented in various forms. When comparing two sets of competing IVs in regression, such as an AR model versus a straight line (first-order polynomial) model, each competing model's set of IVs is initially evaluated separately in a crude regression model. Subsequently, both sets of IVs are included in an adjusted regression model. The IVs demonstrating similar regression coefficients across both the crude and adjusted models are considered true predictors of the DV. Conversely, significant discrepancies between the crude and adjusted IV estimators indicate confounding variables. This method relies on a comparative assessment of estimators rather than a formal statistical test.

Table 4
Results of Linear Time Trend (IV: ID_Day) + AR(1, 7, 14) Regression Model, Using R With OLS Estimation (Confounding Test)

Parameter	Estimate	*SE*	*t*	Pr(>|*t*|)
Intercept	4.453	2.489	1.789	0.078
ID_Day	0.009	0.015	0.585	0.560
Y_{t-1}	0.024	0.085	0.281	0.780
Y_{t-7}	0.418	0.120	3.471	<0.001
Y_{t-14}	0.362	0.123	2.939	0.005

Note. Dependent variable: Y_t, daily smoking of P48. Compare with Table 1 to evaluate whether the addition of the *ID_Day* IV, time trend, modifies the AR coefficients. IV = independent variable; AR = autoregressive; OLS = ordinary least squares.

Given that the crude models have already been calculated and presented in Table 1 for the AR(1, 7, 14) model and in Table 3 for the straight-line model, Table 4 displays the results for the model integrating both. The results in Table 4 reveal two key aspects: (a) the coefficient for the measurement day (*ID_Day* variable), which was significant in Table 3, loses significance in the confounding test of Table 4, with a considerable decrease in its absolute value (from .041, $p = .010$ in Table 3 to .009, $p = .560$ in Table 4); and (b) the coefficients of the AR(1, 7, 14) model, as well as their significance, remain virtually unchanged. Therefore, we can conclude that the confounding tests confirm that the AR model is more suitable for our longitudinal data than the straight-line (first-order polynomial) model.

Discussion

In this study, we have shown that (a) AR models in TS analysis essentially represent a specialized form of regression, with (b) the IVs being prior values of the DV itself, and so (c) since AR models are a special type of linear regression, it is not necessary to use a specific program or a TS analysis module; an ordinary regression module is sufficient; it only requires understanding and knowing how to implement the lagging of the series to be analyzed. Identification of the series through ACF and PACF analysis simplifies the implementation of series lags, and interpreting regression coefficients in TS mirrors that of standard linear regression techniques. Utilizing OLS for TS regression yields valid, or BLUE, which are optimal under conditions of residual non-AC, normality, and homoscedasticity, criteria verified in our data.

Our analyses employed two statistical packages prevalent within the quantitative psychology research community: R and SPSS. R offers the advantages of being cost-free and highly flexible. However, its primary reliance on syntax for operation (despite the availability of numerous online tutorials) can pose challenges. A noted limitation with R is its handling of missing data in

forecasted and residual calculations, which necessitates careful line spelling and data review to understand program-imposed deletions. Conversely, SPSS, while user-friendly and robust, incurs costs and lacks some advanced statistical tests without specific modules. Particularly, SPSS's `TSMODEL` module cannot perform temporal regression using OLS estimation. A critical examination revealed that SPSS's parameter estimation with ML estimation within the SPSS `TSMODEL` module leads to inaccuracies, as demonstrated by the discrepancies in expected values in Table 2, thereby its result is a fallacy (Appendix C).

We advocate for (a) overparameterization when facing uncertainty regarding the number of coefficients to estimate in an AR model, and (b) paying attention to possible seasonality in the model according to the data source; in our case, the data are daily, so we hypothesize a weekly, or 7 days, seasonality. An example can be found in Figure 5, where a traditional TS analyst might have suggested estimating an AR(1,7) model based on significant ACF and PACF at those lags. However, we opted to include AR(14) as part of our hypothesis on seasonality of 7 days (weekly). The inclusion of lags 7 and 14 proved significant in the regression model, despite PACF at lag 14 not showing significance. In reality, only the values from an AR(7,14) model were significant, but we retained AR(1) within our hypothesis, even though it did not prove significant.

Emphasizing the crucial role of understanding and implementing the variable lagging process in TS data analysis, followed by linear regression using standard software, is essential. Yet, verifying the model's conditions (ensuring residuals resemble white noise and exhibit normality and homoscedasticity) is paramount. When dealing with TS data, especially in tasks like forecasting or regression analysis, it is vital to ensure that the treatment of lagged variables is in line with both your analytical objectives and the underlying assumptions of the model used. Mastery of these concepts not only aids in navigating TS data analysis but also enhances the analysis of more intricate temporal models, including multivariate TS and panel data models.

In discussing AR models, such as those examined in this study, it is pivotal to acknowledge that the data series itself serves as the DV to be modeled and forecasted. Simultaneously, when these same data values are suitably lagged, they act as the IVs within the model. Essentially, within a single equation, the variable functions both as the DV and the IV.

An important aspect to consider in the estimation of AR models for TS is the assumption of linearity and the presence of white noise residuals. What happens when these conditions are not met? In such cases, it is essential to closely examine whether the relationship between the IVs and DV is linear or if the variance is subject to processes with conditional heteroscedasticity (e.g., AR with conditional heteroscedasticity, or ARCH models). When residuals are not normally distributed, generalized linear models can be utilized for estimation. Additionally, if the relationship between variables is nonlinear, nonparametric, or semiparametric, AR models, wavelet transforms, and other sophisticated techniques should be considered (Huffaker et al., 2017; Tsay & Chen, 2018). These advanced models, however, extend beyond the scope of this tutorial.

Among the array of statistical linear models frequently utilized in data analysis (including ARIMA, exponential smoothing, and others), AR models stand out for their suitability. This preference stems from their ability to align with theoretical frameworks guided by specific hypotheses, offering straightforward interpretation. In contrast, differentiation models, like Box and Jenkins *I* models, could disrupt the concept of homeostasis fundamental to psychology and biology by their very design. Similarly, MA models under the Box and Jenkins framework and exponential smoothing models primarily serve as data-mining mechanisms, lacking a solid foundation in either hypothesis or theory.

In the section "Data analysis 3," a statistical analysis was conducted to fit our TS data using a straight line (order 1 polynomial). It was found that this approach does not eliminate AC, leading to Type I errors, meaning significant results that, when subjected to the confounding variable test, are shown to be nonsignificant.

We have observed that straight-line models, and by extension, any polynomial models, tend to approach values of plus or minus infinity. This characteristic renders them unsuitable as descriptors of everyday human behavior. One might consider piecewise models, in which a new straight-line model is proposed whenever the predicted behavior tends to exceed established reasonable limits. However, this approach is impractical, as a correct statistical model should be applicable throughout its entire defined trajectory. Furthermore, as previously indicated, this would violate the principle of homeostasis in psychology or biology (Goldstein, 2019; Goldstein & McEwen, 2002; Jonas & Stollberg, 2021). Homeostasis refers to the ability of any organism or system to maintain internal stability despite external changes. In the context of human behavior, it implies that individuals strive to maintain a stable internal state, adjusting their actions to counteract disturbances. Therefore, a model that frequently requires adjustments contradicts the inherent stability and self-regulation observed in human behavior.

Regarding the model's suitability, the conditions surrounding residuals (namely, their lack of ACR, normality, and homoscedasticity) are crucial. Particularly paramount is ensuring that error ACF and PACF is nonsignificant, as this safeguards against the risk of Type I errors in parameter estimation. Additionally, concerning the model's stability, various criteria exist, such as stationarity and stability, that can be checked by the L&B, the S&W, and the B&P tests of residuals. However, a practical approach (Engle & Granger, 1987) is to ensure that residuals exhibit white noise characteristics and that the series demonstrates stability through tests like the LTT.

In this research, we utilized the "straight line" to model the functional relationship between the IV time (*ID_Day*) and the DV daily smoking (Y_t, daily smoking of P48). It is important to note that this linear relationship represents a specific instance of a polynomial relationship. Squaring the time variable would yield a second-order polynomial, with higher orders (cubic, quartic, etc.) obtainable through further exponentiation. The scientific literature presents similar studies comparing second-order polynomial models with AR models (Rosel et al., 2019). These studies consistently demonstrate that AR models provide a superior theoretical and statistical fit. Specifically, the second-order polynomial model failed to eliminate ACR, resulting in residuals that did not exhibit white noise (despite significant global fit parameters and coefficients). Furthermore, the LTT in the polynomial model yielded an undefined value ($-\infty$), whereas the AR model produced a valid LTT statistic. Finally, the confounding variables test confirmed the superiority of the AR model over the second-order polynomial model.

As advancements in technology facilitate the collection of longitudinal data across biological and psychological domains, it becomes increasingly important to adopt analytical methods suited to the structural complexities of temporal data. Consequently,

there is a pressing need to equip upcoming researchers with proficiency in longitudinal data analysis techniques, including TS, multilevel regression, SEM, and latent curve modeling. Furthermore, promoting the open access publication of data sets in scientific journals would enhance their reanalysis and replicability. And finally, in any research with longitudinal data, it would be necessary to request verification that the residuals are "white noise." The aforementioned aspects would favor the advancement of psychology as a science.

References

Aitken, A. C. (1936). IV.—On least squares and linear combination of observations. *Proceedings of the Royal Society of Edinburgh*, *55*, 42–48. https://doi.org/10.1017/S0370164600014346

Andersen, H. K. (2022). A closer look at random and fixed effects panel regression in structural equation modeling using lavaan. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(3), 476–486. https://doi.org/10.1080/10705511.2021.1963255

Anderson, T. W. (2011). *The statistical analysis of time series*. John Wiley & Sons.

Antoniou, T., Martins, D., Campbell, T., Tadrous, M., Munro, C., Leece, P., Mamdani, M., Juurlink, D. N., & Gomes, T. (2021). Impact of policy changes on the provision of naloxone by pharmacies in Ontario, Canada: A population-based time–series analysis. *Addiction*, *116*(6), 1514–1520. https://doi.org/10.1111/add.15324

Ariens, S., Ceulemans, E., & Adolf, J. K. (2020). Time series analysis of intensive longitudinal data in psychosomatic research: A methodological overview. *Journal of Psychosomatic Research*, *137*, Article 110191. https://doi.org/10.1016/j.jpsychores.2020.110191

Arnau, J., & Bono, R. (2001). Autocorrelation and bias in short time series: An alternative estimator. *Quality and Quantity*, *35*(5), 365–387. https://doi.org/10.1023/A:1012223430234

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388. https://doi.org/10.1080/10705511.2017.1406803

Baek, E. K., & Ferron, J. M. (2013). Multilevel models for multiple-baseline data: Modeling across-participant variation in autocorrelation and residual variance. *Behavior Research Methods*, *45*(1), 65–74. https://doi.org/10.3758/s13428-012-0231-z

Basu, D. (2020). Bias of OLS estimators due to exclusion of relevant variables and inclusion of irrelevant variables. *Oxford Bulletin of Economics and Statistics*, *82*(1), 209–234. https://doi.org/10.1111/obes.12322

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*(1), 579–616. https://doi.org/10.1146/annurev.psych.54.101601.145030

Bollen, K. A., & Brand, J. E. (2010). A general panel model with random and fixed effects: A structural equations approach. *Social Forces*, *89*(1), 1–34. https://doi.org/10.1353/sof.2010.0072

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., van Borkulo, C. D., van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), Article 58. https://doi.org/10.1038/s43586-021-00055-w

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. Holden-Day.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time series analysis: Forecasting and control* (5th ed.). John Wiley & Sons.

Box-Steffensmeier, J. M., Freeman, J. R., Pevehouse, J. C., & Hitt, M. P. (2014). *Time series analysis for the social sciences*. Cambridge University Press.

Bringmann, L. F. (2021). Person-specific networks in psychopathology: Past, present, and future. *Current Opinion in Psychology*, *41*, 59–64. https://doi.org/10.1016/j.copsyc.2021.03.004

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE*, *8*(4), Article e60188. https://doi.org/10.1371/journal.pone.0060188

Browne, M. W., & Nesselroade, J. R. (2005). Representing psychological processes with dynamic factor models: Some promising uses and extensions of ARMA time series models'. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Advances in psychometrics: A festschrift for Roderick P. McDonald* (pp. 415–452). L. Erlbaum.

Bye, E. K. (2007). Alcohol and violence: Use of possible confounders in a time-series analysis. *Addiction*, *102*(3), 369–376. https://doi.org/10.1111/j.1360-0443.2006.01701.x

Campos, J., Ericsson, N. R., & Hendry, D. F. (2005). General-to-Specific modeling: An overview and selected bibliography. *The Federal Reserve Board, Paper No. 838*. https://doi.org/10.2139/ssrn.791684

Chowdhury, S. D., & Wang, E. Z. (2009). Institutional activism types and CEO compensation: A time-series analysis of large Canadian corporations. *Journal of Management*, *35*(1), 5–36. https://doi.org/10.1177/0149206308326772

Cochrane, D., & Orcutt, G. H. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *Journal of the American Statistical Association*, *44*(245), 32–61. https://doi.org/10.1080/01621459.1949.10483290

Cundiff, J. M., & Matthews, K. A. (2018). Friends with health benefits: The long-term benefits of early peer social integration for blood pressure and obesity in midlife. *Psychological Science*, *29*(5), 814–823. https://doi.org/10.1177/0956797617746510

Dixon, W. J. (Ed.). (1990). *BMDP statistical software* (Vol. 1). University of California Press.

Dobrow, S. R., Ganzach, Y., & Liu, Y. (2018). Time and job satisfaction: A longitudinal study of the differential roles of age and tenure. *Journal of Management*, *44*(7), 2558–2579. https://doi.org/10.1177/0149206315624962

Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251–276. https://doi.org/10.2307/1913236

Ferrer, E., & Zhang, G. (2009). Time series models for examining psychological processes: Applications and new developments. In R. E. Millsap & A. Maydeu-Olivares (Eds.), *Handbook of quantitative methods in psychology* (pp. 637–657). Sage Publications.

Ferron, J. M., Bell, B. A., Hess, M. R., Rendina-Gobioff, G., & Hibbard, S. T. (2009). Making treatment effect inferences from multiple-baseline data: The utility of multilevel modeling approaches. *Behavior Research Methods*, *41*(2), 372–84. https://doi.org/10.3758/BRM.41.2.372

Fisher, A. J., & Boswell, J. F. (2016). Enhancing the personalization of psychotherapy with dynamic assessment and modeling. *Assessment*, *23*(4), 496–506. https://doi.org/10.1177/1073191116638735

Flor-Arasil, P., Rosel, J. F., Ferrer, E., Barrós-Loscertale, A., & Machancoses, F. H. (2021). Longitudinal effects of distress and its management during COVID-19 lockdown in Spain. *Frontiers in Psychology*, *12*, Article 772040. https://doi.org/10.3389/fpsyg.2021.772040

Florens, J.-P., & Mouchart, M. (1985). Conditioning in dynamic models. *Journal of Time Series Analysis*, *6*(1), 15–34. https://doi.org/10.1111/j.1467-9892.1985.tb00395.x

Freedman, D. A. (2012a). *Statistical models and causal inference*. Cambridge University Press.

Freedman, D. A. (2012b). *Statistical models: Theory and practice* (2nd ed.). Cambridge University Press.

Fried, E. I., Papanikolaou, F., & Epskamp, S. (2022). Mental health and social contact during the COVID-19 pandemic: An ecological momentary assessment study. *Clinical Psychological Science*, *10*(2), 340–354. https://doi.org/10.1177/21677026211017839

Goldstein, D. S. (2019). How does homeostasis happen? Integrative physiological, systems biological, and evolutionary perspectives. *American Journal of*

Physiology-Regulatory, Integrative and Comparative Physiology, *316*(4), R301–R317. https://doi.org/10.1152/ajpregu.00396.2018

Goldstein, D. S., & McEwen, B. (2002). Allostasis, homeostats, and the nature of stress. *Stress*, *5*(1), 55–58. https://doi.org/10.1080/10253890290001234 5

Greene, W. H. (2012). *Econometric analysis* (8th ed.). Pearson.

Hamaker, E. L., & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. *Current Directions in Psychological Science*, *26*(1), 10–15. https://doi.org/10.1177/0963721416666518

Hammond, D., & Gast, D. L. (2010). Descriptive analysis of single subject research designs: 1983–2007. *Education and Training in Autism and Developmental Disabilities*, *45*(2), 187–202. https://www.jstor.org/stable/23879806

Hendry, D. F. (2024). A brief history of general-to-specific modelling. *Oxford Bulletin of Economics and Statistics*, *86*(1), 1–20. https://doi.org/10.1111/obes.12578

Hendry, D. F., Pagan, A. R., & Sargan, J. D. (1984). Dynamic specification. In Z. Griliches & M. D. Intriligator (Eds.), *Handbook of econometrics* (Vol. 2, pp. 1023–1100). North Holland. https://doi.org/10.1016/S1573-4412(84)02010-9

Hendry, D. F., & Trivedi, P. K. (1972). Maximum likelihood estimation of difference equations with moving average errors: A simulation study. *The Review of Economic Studies*, *39*(2), 117–145. https://doi.org/10.2307/2296866

Hibbs, D. (1974). Problems of statistical estimation and causal inference in time-series regression models. In H. L. Costner (Ed.), *Sociological methodology* (pp. 252–308). Jossey-Bass.

Hintzman, D. L. (1980). Simpson's paradox and the analysis of memory retrieval. *Psychological Review*, *87*(4), 398–410. https://doi.org/10.1037/0033-295X.87.4.398

Hoeppner, B. B., Goodwin, M. S., Velicer, W. F., & Heltshe, J. (2007). An applied example of pooled time series analysis: Cardiovascular reactivity to stressors in children with autism. *Multivariate Behavioral Research*, *42*(4), 707–727. https://doi.org/10.1080/00273170701755291

Hohm, I., Wormley, A. S., Schaller, M., &Varnum, M. E. W. (2024). Homo temporus: Seasonal cycles as a fundamental source of variation in human psychology. *Perspectives on Psychological Science*, *19*(1), 151–172. https://doi.org/10.1177/17456916231178695

Huckfeldt, R. R., Kohfeld, C. W., & Likens, T. W. (1982). *Dynamic modeling: An introduction*. Sage Publications.

Huffaker, R., Bittelli, M., & Rosa, R. (2017). *Nonlinear time series analysis with R*. Oxford University Press.

Huitema, B. E., Mckean, J. W., & Mcknight, S. (1999). Autocorrelation effects on least-squares intervention analysis of short time series. *Educational and Psychological Measurement*, *59*(5), 767–786. https://doi.org/10.1177/00131649921970134

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts. https://otexts.com/fpp2/

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.

IBM Corp. (2022). *IBM SPSS statistics for Windows, v29*.

Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: Examining and forecasting change. *Frontiers in Psychology*, *6*, Article 727. https://doi.org/10.3389/fpsyg.2015.00727

Jonas, E., & Stollberg, J. (2021). Homeostasis as affective-motivational state: A threat and defense perspective. *Psychological Inquiry*, *32*(4), 247–252. https://doi.org/10.1080/1047840X.2021.2004818

Jones, R. R., Vaught, R. S., & Weinrott, M. (1977). Time-series analysis in operant research. *Journal of Applied Behavior Analysis*, *10*(1), 151–166. https://doi.org/10.1901/jaba.1977.10-151

Kazdin, A. E. (1982). Single-case experimental designs in clinical research and practice. *New Directions for Methodology of Social & Behavioral Science*, *13*, 33–47. https://doi.org/10.1016/j.brat.2018.11.015

Kazdin, A. E. (2019). Single-case experimental designs. Evaluating interventions in research and clinical practice. *Behaviour Research and Therapy*, *117*(1), 3–17. https://doi.org/10.1016/j.brat.2018.11.015

Kazdin, A. E. (2021). Single-case experimental designs: Characteristics, changes, and challenges. *Journal of the Experimental Analysis of Behavior*, *115*(1), 56–85. https://doi.org/10.1002/jeab.638

Kiesner, J., Mendle, J., Eisenlohr-Moul, T. A., & Pastore, M. (2016). Cyclical symptom change across the menstrual cycle. *Clinical Psychological Science*, *4*(5), 882–894. https://doi.org/10.1177/2167702616635031

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*, Article 513. https://doi.org/10.3389/fpsyg.2013.00513

Kitagawa, G. (2021). *Introduction to time series modeling with applications in R* (2nd ed.). CRC, Taylor & Francis.

Kmenta, J. (1971). *Elements of econometrics*. Macmillan.

Kuba, K., & Scheibe, S. (2017). Let it be and keep on going! Acceptance and daily occupational well-being in relation to negative work events. *Journal of Occupational Health Psychology*, *22*(1), 59–70. https://doi.org/10.1037/a0040149

Kumar, A., Koullia, N., Jongenburger, M., Koutris, M., Lobbezoo, F., Trulsson, M., & Svensson, P. (2019). Behavioral learning and skill acquisition during a natural yet novel biting task. *Physiology & Behavior*, *211*, Article 112667. https://doi.org/10.1016/j.physbeh.2019.112667

Lüdtke, O., & Robitzsch, A. (2022). A comparison of different approaches for estimating cross-lagged effects from a causal inference perspective. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(6), 888–907. https://doi.org/10.1080/10705511.2022.2065278

Maican, C., Lixandroiu, R., & Constantin, C. (2016). Interactivia.ro—A study of a gamification framework using zero-cost tools. *Computers in Human Behavior*, *61*, 186–197. https://doi.org/10.1016/j.chb.2016.03.023

Maldonado, G. & Greenland, S. (1993). Interpreting model coefficients when the true model form is unknown. *Epidemiology*, *4*(4), 310–318. https://doi.org/10.1097/00001648-199307000-00006

Markey, P. M., & Markey, C. N. (2013). Seasonal variation in internet keyword searches: A proxy assessment of sex mating behaviors. *Archives of Sexual Behavior*, *42*(4), 515–521. https://doi.org/10.1007/s10508-012-9996-5

McCarthy, D. E., Minami, H., Yeh, V. M., & Bold, K. W. (2015). An experimental investigation of reactivity to ecological momentary assessment frequency among adults trying to quit smoking. *Addiction*, *110*(10), 1549–1560. https://doi.org/10.1111/add.12996

McDonald, S., Quinn, F., Vieira, R., O'Brien, N., White, M., Johnston, D. W., & Sniehotta, F. F. (2017). The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: A systematic literature overview. *Health Psychology Review*, *11*(4), 307–323. https://doi.org/10.1080/17437199.2017.1316672

Meyer, C., Muto, V., Jaspar, M., Kussé, C., Lambot, E., Chellappa, S. L., Degueldre, C., Balteau, E., Luxen, A., Middleton, B., Archer, S. N., Collette, F., Dijk, D.-J., Phillips, C., Maquet, P., & Vandewalle, G. (2016). Seasonality in human cognitive brain responses. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(11), 3066–3071. https://doi.org/10.1073/pnas.1518129113

Molenaar, P. C. M. (1985). A dynamic factor analysis model for the analysis of multivariate time series. *Psychometrika*, *50*(2), 181–202. https://doi.org/10.1007/BF02294246

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology—This time forever. *Measurement*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Munafò, M. R., Hitsman, B., Rende, R., Metcalfe, C., & Niaura, R. (2008). Effects of progression to cigarette smoking on depressed mood in adolescents: Evidence from the National Longitudinal Study of Adolescent

Health. *Addiction*, *103*(1), 162–171. https://doi.org/10.1111/j.1360-0443.2007.02052.x

Murray, S. L., Griffin, D. W., Derrick, J. L., Harris, B., Aloni, M., & Leder, S. (2011). Tempting fate or inviting happiness?: Unrealistic idealization prevents the decline of marital satisfaction. *Psychological Science*, *22*(5), 619–626. https://doi.org/10.1177/0956797611403155

Nestler, S., & Humberg, S. (2024). Univariate autoregressive structural equation models as mixed-effects models. *Structural Equation Modeling: A Multidisciplinary Journal*, *31*(2), 357–366. https://doi.org/10.1080/10705511.2023.2212865

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 703–708. https://doi.org/10.2307/1913610

Nordmann, E., McAleer, P., Toivo, W., Paterson, H., & DeBruine, L. M. (2022). Data visualization using R for researchers who do not use R. *Advances in Methods and Practices in Psychological Science*, *5*(2), Article 25152459221074654. https://doi.org/10.1177/25152459221074654

Panik, M. J. (2014). *Growth curve modeling: Theory and applications*. J. Wiley & Sons.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, *3*(none), 96–146. https://doi.org/10.1214/09-SS057

Pesaran, M. H. (2016). Introduction to dynamic economic modelling. In M. H. Pesaran (Ed.), *Time series and panel data econometrics* (pp. 120–135). Oxford Academic. https://doi.org/10.1093/acprof:oso/9780198736912.003.0006

Powers, J. R., & Young, A. F. (2008). Longitudinal analysis of alcohol consumption and health of middle-aged women in Australia. *Addiction*, *103*(3), 424–432. https://doi.org/10.1111/j.1360-0443.2007.02101.x

R Core Team. (2023a). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

R Core Team. (2023b). *The R stats package V. 4.3.2*. R Foundation for Statistical Computing. https://cran.r-project.org/web/packages/STAT/STAT.pdf

Reilly, D. P. (2019). *AUTOBOX, v7.0*. Automatic Forecasting Systems.

Robinaugh, D. J., Hoekstra, R. H. A., Toner, E. R., & Borsboom, D. (2020). The network approach to psychopathology: A review of the literature 2008-2018 and an agenda for future research. *Psychological Medicine*, *50*(3), 353–366. https://doi.org/10.1017/S0033291719003404

Rosel, J., Arnau, J., & Jara, P. (1998). Relationship between the mean and the constant in a Box–Jenkins time series model. *Quality and Quantity*, *32*(2), 155–163. https://doi.org/10.1023/A:1004200515015

Rosel, J. F., Elipe-Miravet, M., Elósegui, E., Flor-Arasil, P., Machancoses, F. H., Pallarés, J., Puchol, S., & Canales, J. J. (2020). Pooled time series modeling reveals smoking habit memory pattern. *Frontiers in Psychiatry*, *11*, Article 49. https://doi.org/10.3389/fpsyt.2020.00049

Rosel, J. F., & Elósegui, E. (1994). Daily and weekly smoking habits: A Box–Jenkins analysis. *Psychological Reports*, *75*(3_suppl), 1639–1648. https://doi.org/10.2466/pr0.1994.75.3f.1639

Rosel, J. F., Maldonado, E. F., Jara, P., Machancoses, F. H., Pallarés, J., Torrente, P., Puchol, S., & Canales, J. J. (2019). Intensive longitudinal modelling predicts diurnal activity of salivary alpha-amylase. *PLOS ONE*, *14*(1), Article e0209475. https://doi.org/10.1371/journal.pone.0225745

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177. https://doi.org/10.1037/1082-989X.7.2.147

Segerstrom, S. C., Sephton, S. E., & Westgate, P. M. (2017). Intraindividual variability in cortisol: Approaches, illustrations, and recommendations. *Psychoneuroendocrinology*, *78*, 114–124. https://doi.org/10.1016/j.psyneuen.2017.01.026

Shadish, W. R., & Sullivan, K. J. (2011). Characteristics of single-case designs used to assess intervention effects in 2008. *Behavior Research Methods*, *43*(4), 971–980. https://doi.org/10.3758/s13428-011-0111-y

Shapiro, M. B. (1961). The single case in fundamental clinical psychological research. *British Journal of Medical Psychology*, *34*(3–4), 255–262. https://doi.org/10.1111/j.2044-8341.1961.tb00949.x

Shin, Y. (2017). *Time series analysis in the social sciences: The fundamentals*. University of California Press.

Shumway, R. H., & Stoffer, D. S. (2017). *Time series analysis and its applications: With R examples* (4th ed.). Springer.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. Appleton-Century.

Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods*, *17*(4), 510–550. https://doi.org/10.1037/a0029312

Sorokin, P. A. (1964). *Sociocultural causality, space, time: A study of referential principles of sociology and social science*. Russell & Russell.

Sternberg, R. J. (2019). Theories and hypotheses. In R. J. Sternberg (Ed.), *Guide to publishing in psychology journals* (pp. 54–64). Cambridge University Press.

Suen, L., & Ary, D. (1987). Autocorrelation in applied behavior analysis: Myth or reality? *Behavioral Assessment*, *9*(2), 125–130.

Tanious, R., & Onghena, P. (2021). A systematic review of applied single-case research published between 2016 and 2018: Study designs, randomization, data aspects, and data analysis. *Behavior Research Methods*, *53*(4), 1371–1384. https://doi.org/10.3758/s13428-020-01502-4

Tokunaga, H. T. (2021). *Moving from IBM SPSS® to R and RStudio®: A statistics companion*. Sage Publications.

Trapletti, A., Hornik, K., & LeBaron, B. (2024). *Package 'tseries'*. R Foundation for Statistical Computing.

Tsay, R. S., & Chen, R. (2018). *Nonlinear time series analysis*. John Wiley & Sons.

Vandaele, W. (1983). *Applied time series and Box–Jenkins models*. Academic Press.

Velicer, W. F., & Molenaar, P. C. (2013). Time series analysis for psychological research. In I. B. Weiner, J. Schinka, & W. F. Velicer (Eds.), *Handbook of psychology: Vol 2. Research methods in psychology* (2nd ed., pp. 628–660). Wiley.

Velicer, W. F., Redding, C. A., Richmond, R., Greeley, J., & Swift, W. (1992). A time series investigation of three nicotine regulation models. *Addictive Behaviors*, *17*(4), 325–345. https://doi.org/10.1016/0306-4603(92)90039-X

Voelkle, M. C., & Oud, J. H. L. (2013). Continuous time modeling with individually varying time intervals for oscillating and non-oscillating processes. *British Journal of Mathematical and Statistical Psychology*, *66*(1), 103–126. https://doi.org/10.1111/j.2044-8317.2012.02043.x

Walker, G. (1931). On the theory of the linear regression. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *132*(820), 1–20. https://doi.org/10.1098/rspa.1931.0010

Westreich, D. (2020). *Epidemiology by design: A causal approach to the health sciences*. Oxford University Press.

Whittle, P. (1951). *Hypothesis testing in time series analysis* [Doctoral thesis, Uppsala University]. Almqvist & Wiksells.

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis [R package]*. Springer-Verlag. https://ggplot2.tidyverse.org

Wit, E., Heuvel, E., & Romeijn, J. W. (2012). "All models are wrong …": An introduction to model uncertainty. *Statistica Neerlandica*, *66*(3), 217–236. https://doi.org/10.1111/j.1467-9574.2012.00530.x

Wysocki, A. C., Lawson, K. M., & Rhemtulla, M. (2022). Statistical control requires causal justification. *Advances in Methods and Practices in Psychological Science*, *5*(2), Article 25152459221095823. https://doi.org/10.1177/25152459221095823

Yule, G. U. (1903). Notes on the theory of association of attributes in statistics. *Biometrika*, *21*(1), 121–134. https://www.jstor.org/stable/2331677

Yule, G. U. (1927). On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, *226*(267), 298–30. https://doi.org/10.1098/rsta.1927.0007

(*Appendices continue*)

Appendix A

Long-Term Trend (LTT) in Differenced Time Series (TS) Models

This appendix explores the LTT stability of the mean in TS models, particularly focusing on the concept of the LTT as formulated in temporal models (Engle & Granger, 1987; Panik, 2014). LTT is defined under the following conditions:

$$\begin{aligned} &(a)\, t \to \infty, \\ &(b)\, Y_t = Y_{t-1} = Y_{t-2} = \cdots = Y_{t-j} = Y^*, \text{ and} \\ &(c)\, e_t = 0. \end{aligned} \tag{A1}$$

Given these conditions, if Equation 2 posits that the mean of Y_t is not stable and the TS is differenced once, $I(1)$, transforming it into $y_t = Y_t - Y_{t-1}$, then the differenced equation becomes:

$$Y_t - Y_{t-1} = b_0 + b_1(Y_{t-1} - Y_{t-2}) + e'_t. \tag{A2}$$

Applying Equation A1 for LTT:

$$Y^* - Y^* = b_0 + b_1(Y^* - Y^*) + 0, \tag{A3}$$

$$0 = b_0 + b_1(0), \tag{A4}$$

resulting always in $0 = b_0$ (?), but normally $b_0 \neq 0$. This outcome, described in algebraic terms, is identified as a fallacy or a contradiction in identity, indicating that the initial assumptions leading to this conclusion are not tenable. This simple demonstration can be generalized to any higher level of differentiation, illustrating the methodological implications of assuming a nonstable mean in TS models.

Appendix B

LTT in Table 1 Results

This appendix elucidates the LTT derived from the model presented in Table 1 of our study:

$$\begin{aligned} &Y_t = b_0 + b_1 Y_{t-1} + b_2 Y_{t-7} + b_3 Y_{t-14} + e_t, \\ &Y_t = 4.368 + .035 Y_{t-1} + .430 Y_{t-7} + .361 Y_{t-14} + e_t. \end{aligned} \tag{B1}$$

Applying Equations A1 (from Appendix A) to B1, we have:

$$Y^* = 4.368 + .035Y^* + .430Y^* + .361Y^* + 0, \tag{B2}$$

and solving for Y^*: $Y^* - .035Y^* - .430Y^* - .361Y^* = 4.368$,

$$\begin{aligned} &Y^*(1 - .035 - .430 - .361) = 4.368, \quad .174Y^* = 4.368, \\ &Y^* = 4.368/.174. \end{aligned} \tag{B3}$$

Hence, the LTT of the smoking model is $Y^* = 25.103$. To determine if this value is acceptable, considering the mean (M) is 23.19 with a standard deviation of 3.50, the 95% confidence interval (CI) for these values would be $M \pm 1.96\ SD$, or 95% CI = (16.33; 30.05). Therefore, the LTT value of 25.103 falls within an acceptable range. This validates the model outlined in Table 1 and Equation B1 as an accurate representation of the daily smoking pattern for participant P48, so the value of the LTT of the smoking model is: $Y^* = 25.103$. We can ask ourselves if it is an acceptable value, if we know that $M = 23.19$, and the $SD = 3.50$, the 95% confident interval (or 95% CI) of the values will be: $M \pm 1.96\ SD$, or 95% CI = (16.33; 30.05), so the LTT of 25.103 is an acceptable value, and we can undertake the estimated model on Table 1 and Equation B1 as a correct representation of the daily smoking for P48.

Appendix C

Error in Table 2 Equation, SPSS `TSMODEL`

This appendix addresses the error identified in the equation presented in Table 2, which arises from the SPSS `TSMODEL` analysis. First, we will determine (a) the expected values of the variables in Table 2; (b) then, we will isolate the intercept b_0; and (c) finally, we will manually estimate the predicted values of the DV, checking if the mean is not equal to the value zero. We will observe that these predicted values are incorrect due to the error in the constant in Table 2.

a. The equation of Table 2, with `TSMODEL` SPSS, is:

$$Y_t^{\text{SPSS}} = 23.179 + .061 Y_{t-1} + .465 Y_{t-7} + .329 Y_{t-14} + e_t^{\text{SPSS}}. \tag{C1}$$

The expected values in any linear model, including Equation B1, are calculated as follows:

$$E(Y_t) = E(b_0 + b_1 Y_{t-1} + b_2 Y_{t-7} + b_3 Y_{t-14} + e_t), \tag{C2}$$

where $E()$ means each respective expected value, so $E(Y_t) = \bar{Y}$, $Ek = k$, $E(kY) = k\bar{Y}$, and $E(e) = 0$, with Y representing any variable, k any constant (coefficient), and e the set of forecasting errors. Thus, Equation C2 simplifies to:

$$\bar{Y}_t = b_0 + b_1 \bar{Y}_{t-1} + b_2 \bar{Y}_{t-7} + b_3 \bar{Y}_{t-14}, \tag{C3}$$

given the means $\bar{Y}_t = 23.193$, $\bar{Y}_{t-1} = 23.183$, $\bar{Y}_{t-7} = 23.132$,

(*Appendices continue*)

$\bar{Y}_{t-14} = 22.870$, and substituting the coefficients from Table 2 into Equation C3:

$$23.193 = 23.179 + .061 \times 23.183 + .465 \times 23.132 + .329 \times 22.870 = 42.874. \tag{C4}$$

This results in a fallacy: $23.193 = 42.874$ (?), indicating estimation errors in Table 2's TSMODEL SPSS results. Observe that with the SPSS TSMODEL, the values of $b_0 = 23.164$, and of $\bar{Y}_t = 23.193$ are practically the same, except for decimal rounding errors.

Conversely, applying regression coefficients from R, or from REGRESSION SPSS (Table 1 and Equation 5), yields a mean that aligns with the DV:

$$\begin{aligned}\bar{Y}_t &= 4.368 + .035\bar{Y}_{t-1} + .430\bar{Y}_{t-7} + .361\bar{Y}_{t-14} \\ &= 4.368 + .035 \times 23.183 + .430 \times 23.132 \\ &\quad + .361 \times 22.870 \sim 23.193,\end{aligned} \tag{C5}$$

which is the real mean of the DV, $\bar{Y}_t = 23.193$.

The discrepancy in Equation C4 ($23.193 \neq 42.874$) may stem from the methodological approach in maximum likelihood estimation programs, where differencing against the mean of each variable is an initial step to zero-center data, focusing parameter estimation under residual normality. For example, as the Biomedical Program (BMDP) says, "the constant corresponds to the intercept term (usually the mean of the series)" (Dixon, 1990, p. 474), interpreting that the mean and the constant are identical values, when in reality we have verified that they are not. This error was elucidated by Rosel et al. (1998) as the confusion between the mean and the constant in TS models, and, surprisingly, the original error continues to be repeated many years later. In brief, this appendix has underscored the need for careful consideration of the underlying assumptions and computational procedures in model estimation, particularly in the context of TS analysis using software like TSMODEL SPSS.

(b) In a similar manner, we can calculate the value of the intercept (b_0) in Equation 5, Table 1. If we request to estimate (b_0), as in any regression model, we isolate it:

$$\begin{aligned}b_0 &= \bar{Y}_t - b_1\bar{Y}_{t-1} - b_2\bar{Y}_{t-7} - b_3\bar{Y}_{t-14} \\ &= 23.193 - .035 \times 23.183 - .430 \times 23.132 \\ &\quad - .361 \times 22.870 \sim 4.368,\end{aligned} \tag{C6}$$

that is, the value obtained in Table 1.

(c) If we estimate the predicted values ($\hat{Y}_t^{\text{SPSS}}$) throw the Equation C1, Table 2; and having in account that in Figure 9, on day 19th, with $Y_{19} = 28$ (or Y_t), $Y_{18} = 25$ (or Y_{t-1}), $Y_{12} = 27$ (Y_{t-7}), and $Y_5 = 26$ (Y_{t-14}), the forecasted value $\hat{Y}_{19}^{\text{SPSS}}$ is calculated as follows:

$$\hat{Y}_{19}^{\text{SPSS}} = 23.179 + .061 \times 25 + .465 \times 27 + .329 \times 26 = 45.813. \tag{C7}$$

The forecasted error is then: $\hat{e}_{19}^{\text{SPSS}} = Y_{19} - \hat{Y}_{19}^{\text{SPSS}} = 28 - 45.813 = -17.813$; as shown in "P48_TS_Tbl2.sav."

If we calculate the errors using Equation 5 from Table 1, the mean of the errors is: $\overline{e_t} = 0$. However, if we calculate the mean of the errors from Equation C1 or Table 2, the mean error is: $\bar{e}_t^{\text{SPSS}} = -19.511$. Comparing this mean value to zero yields a t value $= -70.652(68)$, $p < .001$, indicating that $\bar{e}_t^{\text{SPSS}}$ is significantly different from zero.

Received August 3, 2024
Revision received July 6, 2025
Accepted July 23, 2025 ■