Journal of Econometrics and Statistics
Vol. 5, Issue 2, 2025, pp. 185-200

Incorporating Autoregressive and State-space Models for Panel Surveys Estimation

Noam Cohen[a], Orit Marom-Shwarts[a] and Tzahi Makovky[a]

[a] *The Central Bureau of Statistics, Jerusalem, Israel*

ABSTRACT
State-space models have gained increasing popularity in econometric inference for panel surveys due to their flexibility and potential to reduce estimation errors. In this article, we apply state-space models to improve the accuracy of design-based estimates in panel surveys. While trend and seasonal components can be conveniently modeled through the state equation, the rotation pattern inherent to panel designs induces a complex correlation structure among monthly estimates. This correlation structure motivates the integration of an autoregressive (AR) model for the residuals in the observation equation. However, estimating the optimal AR parameters presents a challenge, as the parameters themselves influence the residuals they are intended to model—a self-referential problem. We propose a practical approach using the Yule-Walker equations to estimate AR parameters and introduce tools to assess their reliability. In addition, we offer heuristic insights into the method's performance and present empirical results based on monthly labor force survey data.

KEYWORDS
State-space models, Autoregressive models, Panel data surveys.

1. Introduction

Panel surveys play a central role in both econometric inference and official statistics. Many national statistical agencies rely on them as a cost-effective and practical tool for analyzing economic trends and seasonal patterns. In particular, panel surveys enable the estimation of monthly and yearly changes, making them especially valuable for continuous data collection efforts such as the Labor Force Survey (LFS). The LFS, a key application of panel survey methodology, serves as a critical resource for governments and research institutions to monitor labor market dynamics and inform public policy.

Typically, panel surveys employ a rotating sample design, in which subsets of the population—referred to as panels—are repeatedly surveyed over time according to a predefined rotation scheme. Each panel represents a portion of the target population, and the full monthly sample consists of several panels combined to provide comprehensive coverage. Depending on the design, some panels are retained in subsequent months while others are replaced with new panels representing the same subpopula-

CONTACT Author[a]. Email: Avinoam.cohen@mail.huji.ac.il

Article History
Received : 11 March 2025; Revised : 9 April 2025; Accepted : 18 April 2025; Published : 28 June 2025

To cite this paper

Noam Cohen, Orit Marom-Shwarts & Tzahi Makovky (2025). Incorporating Autoregressive and State-space Models for Panel Surveys Estimation. *Journal of Econometrics and Statistics.* 5(2), 185-200.

tions. For a detailed description of sampling and rotation methods in panel surveys, see Fuller and Rao [8].

When the rotation structure is not explicitly accounted for in the estimation process, the variance of the resulting estimates tends to depend solely on the sample size within a given area and time point. As a result, some geographic areas may contain too few observations to support reliable estimates, limiting their usability or preventing publication altogether. In other cases, even if the sample size is adequate, the estimators may be statistically inefficient.

To address these limitations, the statistical theory of small area estimation offers tools to improve inference by borrowing strength from related areas or time periods. Panel surveys are particularly well-suited for such approaches. When an area is sampled under a rotation scheme, even a small sample size may support effective estimation through temporal borrowing—leveraging information from previous or subsequent time points in which the same area was surveyed. This enables periodic analysis, where time series components such as trend, seasonality, and error structures are modeled to estimate time-varying population parameters. A natural and flexible framework for implementing such analysis is provided by state-space models (SSMs).

State-space models that incorporate time series dynamics through the state equation have been successfully applied in numerous contexts. In this article, we focus specifically on extending the SSM framework by incorporating an autoregressive model for the residuals in the observation equation. When survey data originates from a panel survey with a rotating design, the residuals tend to exhibit serial correlation due to the overlap of sampled units across time. The role of the AR model, therefore, is to capture this correlation structure in a manner that reflects the underlying rotation scheme of the panels.

However, integrating an AR process to model the residuals of the same state-space model in which it is embedded introduces a circular dependency: the AR parameters influence the residuals they are intended to describe. As a result, these parameters are typically not estimated jointly with the other SSM components. Instead, a common approach is to estimate them separately using external methods (see Pfeffermann et al. [17]).

In this work, we propose an alternative strategy that enables the estimation of AR parameters within the SSM framework itself, thereby aligning the AR structure more closely with the modeling context. We further introduce diagnostic tools to evaluate the reliability and stability of these parameter estimates. In addition, we offer heuristic insights to support the rationale for this integrated approach.

An empirical evaluation of the proposed methods, based on data from Israel's Labor Force Survey, demonstrates their strong potential. The method yields significantly lower sampling errors compared to seasonally adjusted design-based estimates.

1.1. *Small Area Estimation Methods*

In the present study, we focus on the estimation of population parameters for small areas—defined as geographic or demographic units with insufficient sample sizes to support traditional design-based estimation methods for producing reliable published estimates. By *design-based*, we refer to classical estimation approaches in which the randomness of the data is attributed to the sampling process, rather than to a probabilistic model. Foundational contributions to design-based survey inference include the works of Horvitz and Thompson [10], and Yates and Grundy [21].

When the survey design incorporates rotating panels with overlapping units across time, it becomes possible—and often advantageous—to exploit data from previous time periods to improve estimation for the current period. Early studies that recognized this potential include Patterson [13], Rao and Graham [19], and Cochran [5] Sections 2.10–2.12.

Over time, alternative approaches to small area estimation have emerged. These methods typically combine a *direct* estimator—derived solely from the data for the area and time of interest—with a *synthetic* estimator that incorporates auxiliary information from other areas or time points. The key methodological distinction between such approaches lies in how the weights are assigned to these two components. Notable developments in this context include the Fay and Herriot [7] model, which is based on random effects, and the Battese et al. [2] and and Pfeffermann and Barnard [15] models, which employ mixed linear models. These models, while effective in many applications, generally do not account for temporal correlation and typically treat data from different time periods as independent.

To address this limitation, time series models have been proposed for small area estimation, enabling the borrowing of information across time within the same area. An effective way to implement this is by formulating the time series model within a state-space framework, allowing for estimation using Kalman filter (KF) algorithms. This approach is briefly discussed by Pfeffermann [14], while detailed derivations of the Kalman filter appear in Kalman [11], and Anderson and Moore [1]. Pfeffermann and Burck [16] introduced a state-space formulation in which the state equation evolves as a stochastic process (specifically, a random walk), yet without modeling the correlation structure of residuals across time.

A notable limitation of such models is their assumption of temporal independence in the observation errors, which may not be held in panel surveys with overlapping units. In contrast, Rao and Yu [20] depart from the state-space framework but introduce a model in which sampling errors across time are explicitly allowed to be correlated through an autoregressive structure. This represents an important step toward recognizing and modeling the dependence induced by panel rotation schemes—a challenge that motivates the methodology proposed in the present work.

Recent studies—including the present work—propose methods that integrate two key ideas: (a) the use of a state-space model (SSM) incorporating time series components, and (b) the allowance for correlation in the sampling errors through an autoregressive (AR) structure.

A notable contribution that combines both elements within a state-space framework is the study by Pfeffermann and Tiller [18]. In their approach, the AR parameters are not estimated internally as part of the SSM. Instead, the method relies on an external estimation of the autocorrelations in the residuals, which are then incorporated into the transition matrix of the state equation. Techniques used for this external estimation include linearization, jackknife repeated replication (JRR), and other methods (see Pfeffermann et al. [17]).

The foundational framework for integrating time series dynamics into state-space models is known in the literature as the Basic Structural Model (BSM). The BSM decomposes the time series into underlying components such as trend and seasonal effects [9]. This model has been widely adopted by official statistical agencies, including the U.S. Bureau of Labor Statistics (BLS) and the Australian Bureau of Statistics.

In this paper, we propose a method for estimating the parameters of the AR component internally within the SSM framework. Our approach employs a recursive algorithm based on the Yule-Walker equations to estimate the autocorrelations directly from the model's residual structure. We further introduce tools to assess the credibility and robustness of the resulting estimates.

2. State-space models with time-series parameters

In this chapter, we present the statistical models used in our analysis, along with the proposed estimation method. The first section introduces the Basic Structural Model, which embeds time series components such as trend and seasonality within a state-space framework. The second section extends the BSM by incorporating an autoregressive model to account for structure in the residuals. We propose a recursive estimation procedure for jointly estimating and evaluating all model parameters. Our variables of interest are labor force characteristics at small-area levels. For illustrative purposes, we focus specifically on the *Employed* variable. In the following chapter, we demonstrate the method using real data from a Labor Force Survey, analyzing both the *Employed* and *Unemployed* populations.

2.1. *The BSM model*

Denote by Y_{ti} the percentage of employed people at time t in the population of a small area $i = 1, \ldots, D$. Let y_{ti} be the direct design-based estimate based upon the observations in the small area. Also, let e_{ti} be the error term such that

$$y_{ti} = Y_{ti} + e_{ti}; \quad E\left(e_{ti}\right) = 0$$

We assume that the errors are independent between the different areas, i.e. $E\left(e_{ti}e_{tj}\right) = 0$ for all $i \neq j$. For this reason henceforth we will focus on a specific small area and omit the index i that represents the area. At this stage we do not assume any particular correlation structure between the sampling errors of different time points of the small area.

We assume that the true percentage of employed people in the population of each small area follows the BSM, independently between areas (Harvey [9]). So estimation in every area is based solely on data from that specific area. For a specific small area the BSM model is defined as

$$y_t = Y_t + e_t, \text{ with } e_t \sim N\left(0, \sigma^2\right) \text{ and } E\left(e_t e_\tau\right) = 0, \ \forall t \neq \tau \tag{1}$$

where

$$\begin{aligned}
Y_t &= L_t + S_t && (2)\\
L_t &= L_{t-1} + R_{t-1} + \eta_{Lt}; \qquad \eta_{Lt} \sim N\left(0, \sigma_L^2\right) && (3)\\
R_t &= R_{t-1} + \eta_{Rt}; \qquad \eta_{Rt} \sim N\left(0, \sigma_R^2\right) && (4)\\
S_t &= -\textstyle\sum_{k=1}^{11} S_{t-k} + \eta_{S,t}; \qquad \eta_{St} \sim N\left(0, \sigma_S^2\right) && (5)
\end{aligned}$$

This model includes thirteen parameters of seasonality and trend for a specific small area. In (1) y_t indicates the direct design-based estimate at time t. It is assumed that the true population size Y_t is defined by the set of equations (1) to (5) of the BSM model. At equation (3) L_t represents a local linear trend of the time series. Equation (4) represents the slope which is a random walk and S_t represents the seasonality. The

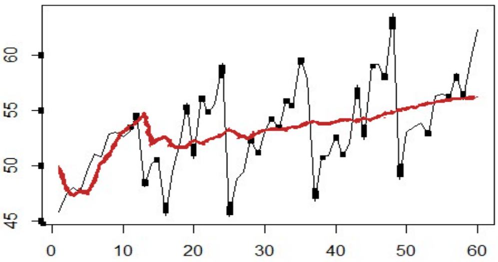

Figure 1. Results of estimating a BSM model on a simulated seasonal series by the use of the KF procedure. The X axis represents time, and the Y axis represents labor force employed proportions with a strong seasonality effect.

seasonality model is broken down into twelve cyclic coefficients. The extra noise η_{St} allows the seasonal terms to progress stochastically across the timeline t. This model is referred to as the additive seasonality model.

A more complex alternative, and marginally more efficient, is the trigonometric seasonal model (e.g. Harvey [9], ch. 2). Nevertheless, we will focus on the additive model which is simpler to interpret with the transition matrix T of the SSM. Finally, all error components $\eta_{Rt}, \eta_{Lt}, \eta_{St}$ are independent white noises.

Let $Z_t = (1,0,1,0,0,0,0,0,0,0,0,0,0)^t$ be an indicator vector and $\beta_t = (L_t, R_t, S_t, S_{t-1}, \ldots, S_{t-10})^t$ be the state vector such that $Z_t\beta_t = L_t + S_t$. Define also the error vector $\eta_t = (\eta_{Lt}, \eta_{Rt}, \eta_{St}, 0, \ldots, 0)^t$ of size 13. We define the 13×13 transition matrix T_{BSM} as follows:

$$T_{BSM} = \begin{bmatrix} 1 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & -1 & -1 & \ldots & -1 & -1 \\ 0 & 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ddots & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & 1 & 0 \end{bmatrix}$$

Hence equations(1) - (5) can be written in the form of a state-space model in the following way (Morris and Pfeffermann [12], eq. 3.4, Pfeffermann [14], eq. 2.1 and others)

$$\begin{aligned} \text{Observation:} \quad & y_t = Z_t\beta_t + e_t \\ \text{State:} \quad & \beta_t = T_{BSM}\beta_{t-1} + \eta_t. \end{aligned}$$

The estimation process can be carried out using several approaches, with the Kalman Filter (KF) being a common and effective choice. Figure 1 illustrates the application of the Kalman Filter to a time series exhibiting seasonality, a positive trend, and a zero-mean error term, with no autocorrelation across time points. The black line represents the original observed series, while the red line displays the seasonally adjusted estimates produced by the Kalman Filter. As shown, the KF yields stable and smooth seasonally adjusted values. Notably, the filtered series undergoes an initial 12-month learning phase, during which the estimates closely follow the original series. After this period, the KF converges to more refined, seasonally adjusted signals.

In the following section we present the effect of combining autoregressive model assumptions for the error terms e_t in (1).

2.2. *The BSM + AR(p) model*

As mentioned in the introduction, the Basic Structural Model can be extended to accommodate correlated model errors across certain time periods. This extension is motivated by the rotation structure of the panel design used in the survey, which creates an overlap between samples from consecutive months and up to several months apart. Such overlap induces correlation in the design-based estimates, and consequently, in the model residuals. Modeling these residuals as an autoregressive process and integrating them within the state-space framework can significantly enhance parameter estimation, beyond the improvement already achieved by the standard BSM.

For example, in the 4-8-4 rotation scheme used in Israel's Labor Force Survey (LFS), each panel is interviewed for four consecutive months, followed by an eight-month gap, and then re-interviewed for another four consecutive months. This design creates overlapping samples with dependencies spanning up to 15 months. Therefore, an appropriate model for the residuals is an AR(15) process, represented as $e_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_{15} e_{t-15} + \eta_{et}$ where η_{et} is a normal random error with zero expectation and standard deviation σ_e^2.

To achieve the incorporation of the autoregressive component into the model structure, the AR coefficients θ_k are embedded into the matrix T, and the lagged residuals e_{t-k} are included as additional state components. This formulation results in a compact and internally consistent representation, especially when compared with alternative approaches found in the literature.

Let the augmented state vector at time t be defined as

$$\beta_t = (L_t, R_t, S_t, S_{t-1}, \ldots, S_{t-10}, e_t, e_{t-1}, \ldots, e_{t-p+1})^t .$$

The indicator vector Z_t will accordingly assign a weight of 1 to the components contributing to the observed series, namely L_t, R_t and e_t, and *0* elsewhere. So $Z_t \beta_t = L_t + R_t + e_t$. The error vector of the state equation is extended accordingly

$$\eta_t = (\eta_{Lt}, \eta_{Rt}, \eta_{St}, 0, \ldots, 0,\ \eta_{et},\ 0,\ \ldots, 0)^t$$

. which now also includes the error component η_{et} resulting from the error in the autoregressive model. Finally, the transition matrix $T_{BSM+ARp}$ is augmented to include the AR coefficients, i.e.,

$$T_{BSM+AR(p)} = \begin{bmatrix} T_{BSM} & 0 \\ 0 & T_e \end{bmatrix}$$

where

$$T_e = \begin{bmatrix} \theta_1 & \theta_2 & \ldots & \theta_p \\ 1 & 0 & \ldots & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} .$$

Since the variables from the AR model are inserted into the state vector β_t, then the observation equation is now $y_t = Z_t\beta_t + \epsilon_t$, hoping that a sound signal in the AR model will substantially reduce the new error term ϵ_t. This unified state-space formulation allows the entire system—including trend, seasonality, and autocorrelated noise—to be estimated coherently within a single recursive framework.

2.3. *Fixed-point estimation of AR coefficients*

The parameters $\theta_1, \theta_2, \ldots, \theta_p$ are the regression coefficients in the AR model $e_t = \sum_{k=1}^{p} \theta_k e_{t-k} + \eta_{et}$ with $\eta_{et} \sim N\left(0, \sigma_e^2\right)$.

A natural and commonly used method for estimating these coefficients is the Yule-Walker equations[1], which relate the autocorrelations of a time series to the parameters of its autoregressive representation. However, in the context of a BSM + AR(p) model, the direct use of the Yule-Walker equations introduces a logical circularity due to a self-referenced situation. Specifically, to apply the Yule-Walker method, one must first compute the autocorrelation structure of the residual series $e_1, \ldots, e_t$. But these residuals are themselves a function of the AR coefficients: they are derived from the state-space model which already includes the parameters coefficients $\theta_1, \theta_2, \ldots, \theta_p$. Thus, we are faced with a self-referential estimation problem: we seek the very coefficients that, when used within the BSM + AR model, produce a residual series that conforms to an AR process governed by those same coefficients.

Formally, we define a mapping f that takes an initial set of AR coefficients $\theta^{(k)} = (\theta_1^{(k)}, \ldots, \theta_p^{(k)})$ and returns a new set $\theta^{(k+1)}$ based on the Yule-Walker solution applied to the residuals produced by the BSM + AR(p) model using $\theta^{(k)}$. We are therefore seeking a fixed point of the mapping f, that is $\theta^{(*)} = f(\theta^{(*)})$, where $\theta^{(*)}$is the set of AR coefficients that remains invariant under the transformation, meaning that it reproduces itself through the recursive estimation process. Finding such a fixed point provides an internally consistent estimate of the AR parameters from within the structure of the model itself, without relying on external residual diagnostics.

We now present a recursive algorithm for estimating the autoregressive coefficients and suggest several heuristic methods to assess their credibility. The following iterative procedure is designed to converge to the fixed point of the coefficient vector. At the k'th iteration, given the current estimates of the AR coefficients $\widehat{\theta}_1^{(k)}, \widehat{\theta}_2^{(k)}, \ldots, \widehat{\theta}_p^{(k)}$, perform the following steps:

1. Execute a BSM + AR(p) procedure that embeds the values $\widehat{\theta}_i(k)$ into the transition matrix T_e and compute the residuals:

$$e_t(k) = y_t - \widehat{y}_t(k) = y_t - \widehat{L}_t(k) - \widehat{S}_t(k); \text{ for } t = 1 \ldots T$$

   where $\widehat{L}_t(k)$ and $\widehat{S}_t(k)$ denote the estimated trend and seasonal components at iteration k.
2. Compute the autocovariances of the residual series $e_1(k), \ldots, e_T(k)$ up to lag $p-1$:

$$Cov\left(e_t(k), e_{t-l}(k)\right), \quad l = 1 \ldots p$$

3. Apply the Yule-Walker equations using these autocovariances to derive updated

[1] Estimated in R using the function *yw()* with the package *itsmr*

estimates $\widehat{\theta}_i^{(k+1)},\ i = 1 \dots p$.

4. Repeat the procedure until convergence is achieved:

$$\left(\widehat{\theta}_1^{(k)}, \widehat{\theta}_2^{(k)}, \dots, \widehat{\theta}_p^{(k)}\right) \overset{k\to\infty}{\to} (\theta_1, \theta_2, \dots, \theta_p)$$

.

To initialize the algorithm, we begin by applying a Basic Structural Model to the design-based series y_t and compute the initial residuals:

$$e_t(0) = y_t - \widehat{y}_t = y_t - \widehat{L}_t - \widehat{S}_t.$$

Indeed when the series of design based estimates y_t vary far beyond the magnitude of the expected $L_t + S_t$ then $Cov\,(y_t - L_t - S_t,\ y_{t-k} - L_{t-k} - S_{t-k}) \approx Cov\,(y_t,\ y_{t-k})$ which implies that the initial autocorrelations required for the AR coefficient estimation can be reasonably approximated using the design-based series y_t directly, rather than the residuals $e_t(0)$.

2.4. *Theoretical Considerations on Convergence and Uniqueness*

The recursive algorithm proposed for estimating the AR coefficients is structured as a fixed-point iteration: starting from an initial AR coefficient vector $\theta^{(0)}$, each iteration involves embedding $\theta^{(k)}$ into the state-space model, extracting the resulting residual series, and re-estimating the AR parameters via the Yule-Walker equations to obtain $\theta^{(k+1)}$. This defines a mapping $T : \theta \to \theta'$ and the procedure seeks a fixed point θ^* such that $T(\theta^*) = \theta^*$.

Convergence of this fixed-point iteration can be established by proving that the mapping T is a contraction under some norm, i.e., that there exists a constant $0 < \kappa < 1$ such that for any two AR coefficient vectors θ and Ψ,

$$\|T(\theta) - T(\Psi)\| \leq \kappa \|\theta - \Psi\|.$$

Under such a condition, the Banach Fixed Point Theorem guarantees both the existence and uniqueness of the fixed point θ^*, as well as geometric convergence of the iterative procedure to θ^* from any starting point (Bertsekas [3], Chapter 1.4). While a full proof of contraction is challenging due to the nonlinear nature of the Kalman filter step and the dependence of the residuals on θ, we conjecture that the mapping behaves approximately as a contraction near the solution under mild conditions—such as the stability of the Kalman filter, the stationarity of the AR process (roots inside the unit circle), and bounded innovation variance.

As such, we believe that local convergence of the algorithm may hold in general, and that global convergence may occur when initialization is sufficiently close to the fixed point. Regarding uniqueness, although the nonlinear structure of the residual surface theoretically allows for multiple fixed points, our empirical results (Section 4.1) demonstrate fast and stable convergence across different small-area series and initial values. These observations suggest that even in the absence of formal global guarantees, the algorithm consistently converges to a meaningful and robust solution. Ultimately, the quality of the resulting model can be evaluated post-estimation, as demonstrated in the next section.

3. Tests of reliability

In this section we propose several evaluation methods for the reliability of the obtained model as a whole and for the autoregressive part specifically.

3.1. *Agreement between the AR model and panel rotation pattern*

A necessary and straightforward way to assess the proper functioning of the proposed model is to verify its ability to recover pre-specified parameters. To this end, one can simulate data generated from a BSM+AR(p) process and evaluate whether the estimation procedure successfully identifies the true underlying parameters. While such a simulation study provides a technical validation, it offers limited insight beyond algorithmic verification.

However, using panel surveys provides an alternative beyond this simulation study. The panels' rotation structure determines the correlation between the estimates of consecutive months where there is an overlap in the sample. Hence the justification of using an AR(p) on the residues of the BSM. When the AR(p) model is chosen accordingly, we expect a substantial and clear correlation between the BSM+ARp residuals $\widehat{e}_t = y_t - \widehat{L}_t - \widehat{S}_t$ and the estimates of the AR model $\widehat{e}_t = \theta_1 \widehat{e}_{t-1} + \theta_2 \widehat{e}_{t-2} + \ldots + \theta_{p-1} \widehat{e}_{t-p+1}$. This observed correlation should align closely with the known structure of panel overlap, thereby reinforcing both the appropriateness of the model and the validity of the AR component in capturing the serial dependence introduced by the sampling design.

In order to check this phenomenon, we execute the BSM+ARp model using a series of models: AR(1), AR(2), ..., AR(p), AR(p+1). For each l we compute the correlation:

$$C_l \equiv Cov\left(\theta_1 \widehat{e}_{t-1} + \theta_2 \widehat{e}_{t-2} + \ldots + \theta_l \widehat{e}_{t-l},\ y_t - \widehat{L}_t - \widehat{S}_t | AR(l)\right).$$

We expect the following relations

- $C_{l+1} > C_l$ for each l for which there is an overlap between the samples;
- $C_{l+1} \cong C_l$ between non overlapping months;
- and, $C_l \cong C_p$ for all $l > p$.

3.2. *Forecasting and predicting the original series*

A fundamental challenge in assessing the credibility of estimated models lies in the uncertainty regarding the model's validity. This issue becomes even more pronounced in the context of one-time or aperiodic surveys, where standard validation techniques may be limited. While methods such as cross-validation are commonly applied, their conclusions can be questionable in situations where the data distribution differs from that of the target population. In contrast, longitudinal models, such as state-space models, offer more robust opportunities for validation. One effective and practical approach is to use the model to predict future time points based on earlier observations. This out-of-sample forecasting provides an empirical means of assessing the model's predictive accuracy and, by extension, its structural credibility.

A recent utilization of this method for an AR(p) model can be found in Zhang and Yi [22]. Given the original series $(y_1, \ldots, y_T)$, estimate the model and its parameters up to time point $t < T$ and then predict $\widehat{y}_{t+k} = f(y_1, \ldots, y_t)$ for $1 \leq k < T - t$. Next,

compare the predicted value with the known observation y_{t+k}. Successful prediction over the months indicates that the model does explain the variability of observations in the original series. Therefore it can be assumed that it has also, to some extent, correctly characterized the model components of trend, seasonality and the process of the residuals. This can be considered a longitudinal analogy of the cross validation approach. We would like to believe that a good explanation of the model for y_{t+k} based on the observations $y_1, \ldots, y_t$ can indicate a correct explanation of reality of the model components. Note however, that this method is indeed rather conservative and one should not expect exact equality between $\widehat{y}_{t+k}$ and y_{t+k}. The algorithm for the BSM + AR(p) model is detailed below.
At time point t, given the observations $y_1, \ldots, y_t$:

(1) Execute a BSM + AR(p) model and estimate parameters $\widehat{L}_t$, $\widehat{S}_t$ and the AR(p) model coefficients $\widehat{\theta}_1, \widehat{\theta}_2, \ldots, \widehat{\theta}_p$, and conclude the model for $\widehat{e}_t = \widehat{\theta}_1 \widehat{e}_{t-1} + \widehat{\theta}_2 \widehat{e}_{t-2} + \ldots + \widehat{\theta}_{p-1} \widehat{e}_{t-p+1}$ as described in previous sections.
(2) Based upon these estimates predict future parameters using equations:

$$\begin{aligned} \widehat{L}_{t+1} &= \widehat{L}_t + \widehat{R}_t; \\ \widehat{S}_{t+1} &= -\widehat{S}_t - \widehat{S}_{t-1} - \ldots - \widehat{S}_{t-10}; \\ \widehat{e}_{t+1} &= \widehat{\theta}_1 \widehat{e}_t + \widehat{\theta}_2 \widehat{e}_{t-1} + \ldots + \widehat{\theta}_p \widehat{e}_{t-p+1}. \end{aligned}$$

(3) Estimate future near term observations

$$\widehat{y}_{t+1} = \widehat{L}_{t+1} + \widehat{S}_{t+1} + \widehat{e}_{t+1};$$

etc', and compare with real values y_{t+1} etc'.

Initialization of this assessment can start at time points $t = 13$ or $t = 25$ to allow model consideration of all 12 seasonal effects.

3.3. *Convergence of the Autoregressive parameters*

As discussed in Section 2.3, identifying an autoregressive model for the error component of a state-space model presents an inherently circular, self-referential problem. The proposed algorithm aims to address this challenge based on a necessary condition that must be satisfied for model identification. However, in the absence of a formal proof, convergence of the algorithm cannot be guaranteed. Nevertheless, the algorithm represents a crucial step toward estimating the AR structure embedded within the model. The process that estimates $\widehat{\theta}_1(k), \widehat{\theta}_2(k), \ldots, \widehat{\theta}_p(k)$ for $k = 1, 2, \ldots$ may present a slow or bimodal convergence, in which cases compromising the reliability of the estimation.

An additional and practical form of assessment involves conducting a sensitivity analysis on the estimated AR parameters. Once convergence has been achieved, any substantial deviation from the estimated parameter values should result in a significant deterioration in the performance of the state-space model. In particular, imposing an incorrect AR model should lead to filtered estimates that effectively revert toward the original design-based series, rather than capturing the smoothed latent structure. Observing such a degradation provides strong empirical support for the validity of the Yule-Walker-based AR estimates, indicating that they meaningfully capture the underlying autocorrelation structure of the residuals.

4. Application of the proposed methods on labor force data

The Labor Force Survey (LFS) is a principal survey administered by national statistical offices to continuously monitor labor market trends. In Israel, the LFS is conducted as a panel survey following a 4-8-4 rotation pattern: each panel is interviewed for four consecutive months, then excluded from the sample for eight months, and subsequently re-interviewed for an additional four months. This design results in a 75% monthly overlap and a 50% annual overlap among sampled units. The estimation method used in Israel's LFS is known as Composite Estimation (Fuller and Rao, 2001), a refined design-based approach that combines current observations with past data using weighted averages (see Cochran [4], Chapter 7). This method achieves lower sampling variances compared to standard design-based estimators that rely solely on current-period data. For a detailed description of the Israeli LFS and its methodology, see Cohen et al. [6].

Despite its advantages, the sampling errors of the composite estimates remain too high to support the reliable publication of labor force estimates at small-area levels. This issue is especially pronounced in localities with populations between 20,000 and 50,000—the focus of our current study. These areas are sampled with certainty each month, yet the number of observations per month is insufficient for the publication of direct monthly estimates. Our analysis focuses on a five-year period spanning from January 2014 to December 2018. Let $D = 45$ be the number of small Areas and let t be a time index such that $t = 1, 2, \ldots, 60$.

4.1. *Fitting the state-space models*

Fitting a BSM to the time series of employed individuals across various localities resulted in a noticeable reduction in estimate variability for the majority of areas. In approximately two-thirds of the localities, the model demonstrated marked improvement, reflecting its ability to successfully capture underlying trend and seasonality components. We then extended the model by incorporating an autoregressive process for the residuals, resulting in the BSM+AR model. As discussed earlier, the 4-8-4 rotation structure of the Israeli Labor Force Survey (LFS) induces sample overlap across time periods of up to 15 months, making an AR(15) model $e_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_{15} e_{t-15} + \eta_{et}$ a theoretically suitable choice. Embedding this residual structure within the BSM+AR framework further reduced variability in the estimates—even beyond the gains achieved by the BSM alone. Despite the logical appeal of using AR(15), we opted to reduce the number of AR parameters to p=3. The motivation for this simplification is discussed later. As will be shown, even with this limitation, the BSM+AR(3) model achieves comparable efficiency to the full AR(15) model.

Figure 2 presents design-based composite estimates alongside model-based estimates of employment ratios. Each locality is prefixed with the symbol "X." Each of the two graphs displays a time series of 60 monthly estimates, covering five years of data. Notably, the BSM+AR(3) model shows significant improvements, even in localities where the BSM alone was insufficient. This suggests that much of the residual variability stemmed from temporal autocorrelation. For instance, in locality X31, the BSM moderately captured trend and seasonality. In contrast, in locality X874, the

BSM failed to adequately model the underlying structure—the estimated trend line from the Kalman Filter closely tracked the original design-based series. However, the BSM+AR(3) model still achieved notable improvements in this locality, indicating that modeling the residual structure contributed substantially to the overall model performance.

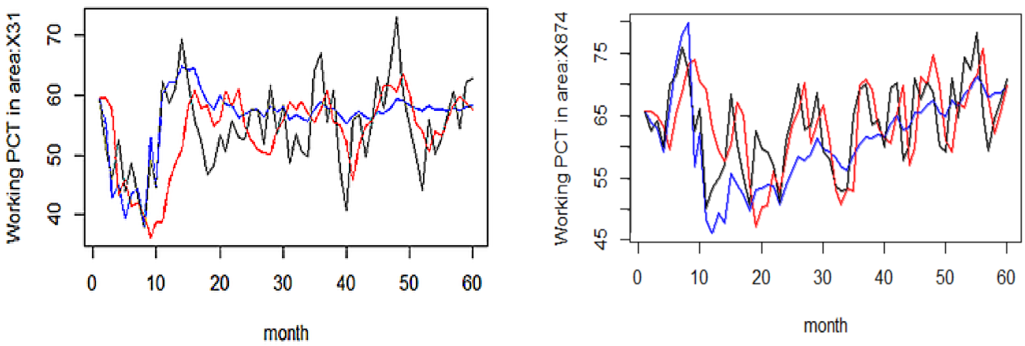

Figure 2. *Design-based composite estimates versus model-based estimates of employment proportions.* The black line represents the design-based estimates, the red line shows the BSM estimates, and the blue line corresponds to the BSM+AR(3) estimates. Locality X874 exemplifies a case where the BSM alone was unable to capture the underlying trend of the time series. However, a clear improvement is observed with the BSM+AR(3) model, indicating its enhanced ability to account for residual autocorrelation.

Table 1 presents a summary of state-space model estimates based on Labor Force Survey (LFS) data for two labor force attributes. Let $\widehat{\nu}_{tk}$ denote the model-based employment estimate estimates for area $k = 1, \ldots, 45$ and $t = 1, \ldots, 60$ time units. For each area k, we define the temporal average of the estimates by $\widehat{\nu}_k$, and by $\widehat{\nu}$ the overall mean across all areas. Similarly, the standard deviation of the estimates across all regions is denoted by $\widehat{\phi}$, and the averaged CV over all areas is given by $\widehat{\phi}/\widehat{\nu}$.

Table 1. Comparison between Composite (design-based), BSM, and BSM+AR(3) for labour force attributes.

		1–60			25–60		
Series	Stat.	Comp.	BSM	BSM+AR(3)	Comp.	BSM	BSM+AR(3)
Employed	Avg	59.48	60.31	60.33	60.72	60.74	60.66
	SD	5.91	5.01	4.53	5.73	4.19	2.00
	CV	10.0	8.3	7.6	9.4	6.9	3.3
Unemployed	Avg	4.39	4.38	4.33	3.94	3.83	3.78
	SD	2.80	1.98	2.04	2.56	1.22	0.68
	CV	64.0	45.4	47.2	65.0	32.0	18.1

As can be observed, for both labor force attributes—employed and unemployed—the design-based (Composite) estimates and the two model-based estimates are broadly consistent in terms of their average values. However, both model-based approaches yield a notable reduction in the variability of the estimates compared to the design-based method. Among them, the BSM+AR(3) model, which accounts for the rotation structure of the survey, provides a substantial improvement in the estimation of both employment and unemployment levels, exceeding the gains achieved by the BSM alone.

In both cases, the Kalman Filter requires an initialization period to learn the seasonal structure of the series, typically at least 12 months. To evaluate the impact of this learning phase, we recalculated the summary statistics starting from time point t=25, corresponding to a two-year learning period. The results aligned with expectations. For instance, in the case of the employment estimates, the standard deviation of design based Composite estimates remained largely unchanged, ranging from 5.73% to 5.91%. In contrast, the standard deviation of the BSM estimates decreased from 5.01% to 4.19%, reflecting an improvement of approximately 20%. The BSM+AR(3) model exhibited a similar stabilization pattern, further reinforcing its effectiveness following the learning phase. The greatest gains were observed in the unemployment estimates, where the model-based methods achieved a reduction of over 50% in the coefficient of variation (CV), indicating a dramatic increase in estimate precision.

4.2. *Assessing the reliability of the models*

4.2.1. Assessment of the autoregressive model

In this section, we present the results of the reliability tests proposed in Chapter 3.

As previously discussed, the rotation structure of the survey panels justifies modeling the residuals of the BSM using an AR(15) process. However, we opted to restrict the model to an AR(3) specification, based on two key considerations:

1. Avoiding overfitting: Given the relatively short time series—60 monthly observations per area—fitting an AR(15) model with 15 parameters poses a risk of overfitting.
2. Marginal gains beyond lag 3: As demonstrated below, the incremental benefit of using more than three lags is limited.

Table 2 below reports the values of the correlation statistic C_l (as defined in Section 3.1) for various lag of l. The results show that the correlation reaches 0.608 for AR(3), and only increases marginally to 0.735 for AR(15), despite the addition of 12 more parameters.

The pattern closely reflects the 4-8-4 panel rotation scheme:

- A sharp increase in C_l is observed from $l = 1$ to $l = 3$, where the overlap between monthly samples is strongest.
- For $l = 4$ to $l = 11$, where no overlap exists, the correlation levels off, i.e., $C_{l+1} \cong C_l$.
- A renewed increase in correlation occurs at $l = 12$ to $l = 15$, coinciding with the reappearance of previously sampled panels, and then stabilizes again.

Table 2. $\rho(q)$ values comparing AR(q) residual structures.

AR(q)	1	2	3	4	8	11	12	15	16
$\rho(q)$	0.390	0.465	**0.603**	0.603	0.629	0.642	0.670	**0.735**	0.735

This empirical behavior fully aligns with the theoretical structure of the 4-8-4 rotation, validating both the modeling approach and the choice of reduced lag order. The use of an AR(3) model therefore captures nearly all meaningful temporal correlation with a minimal number of parameters, offering a favorable trade-off between parsimony and explanatory power.

4.2.2. Assessment of model prediction

The analysis described in Section 3.2 aims to evaluate the model's capacity to decompose the original longitudinal series of design-based estimates into meaningful state-space components. Specifically, if the model can successfully predict future values of the design-based estimate y_{t+1} using only past information $(y_1, \ldots, y_t)$, this reinforces our confidence in the validity and reliability of the model structure.

Following the prediction procedure outlined in Section 3.2, for each time point t, we computed the model parameters $\widehat{L}_{t+1}, \widehat{S}_{t+1}$and $\widehat{e}_{t+1}$ using data of previous time points. These were then used to generate prediction two time points ahead $\widehat{y}_{t+1}, \widehat{y}_{t+2}$. So the prediction $\widehat{y}_{t+k}$ does not rely on y_{t+k}.

The comparison between the predicted series and the original design-based estimates revealed a high level of agreement. Both one-month and two-month-ahead forecasts demonstrated a good accordance with the original series, indicating strong predictive performance. This provides compelling justification for using the model not only for nowcasting current labor force conditions but also for forecasting short-term labor market trends. Of course, this level of predictive accuracy was not observed uniformly across all series—in certain localities, where the noise in the design-based estimates was particularly high, the model struggled to produce reliable predictions. Nevertheless, the overall performance across the majority of areas supports the utility of the modeling approach.

Two most successful examples illustrating this predictive performance are shown in the following figures.

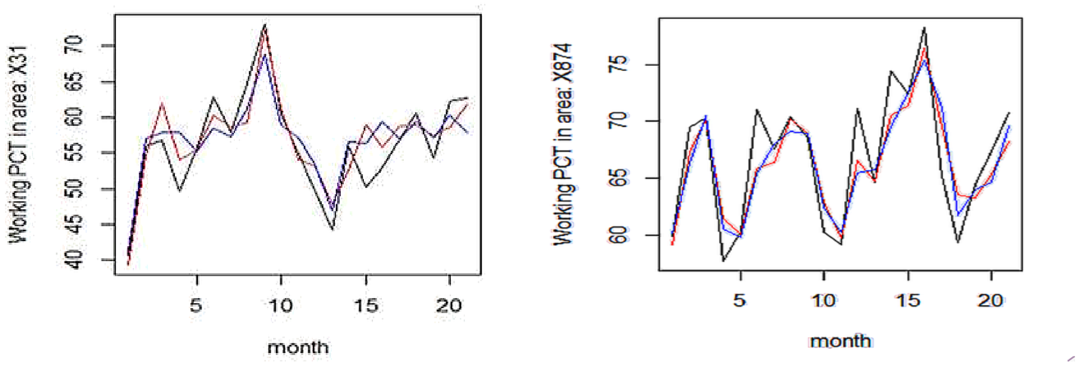

Figure 3. Prediction of the state-space model BSM+AR(3). In the graphs below, black - *original* estimates, red - prediction for $\widehat{y}_{t+1}$ based on observations $y_1, \ldots, y_t$, blue - prediction for $\widehat{y}_{t+2}$ based on observations $y_1, \ldots, y_t$

4.2.3. Assessment of convergence of AR parameters

A plausible method for evaluating the reliability of the estimated autoregressive parameters is to perform a sensitivity analysis. Once convergence of the recursive algorithm described in Section 2.3 has been achieved, we artificially perturbed the AR parameter estimates by up to 25% from their final values. This modification resulted in a substantial degradation of the model's performance: the trend component collapsed, and the model failed to properly estimate the remaining components of the BSM+AR(3) structure. As a result, the filtered series effectively reverted to the original design-based composite estimates. This behavior strongly supports the conclusion that the Yule-Walker estimates accurately capture the true underlying autocorrelation structure in the residuals.

Table 3. Convergence of AR(3) parameters across iterations.

Iteration	ϕ_1	ϕ_2	ϕ_3
$k = 1$	0.4302	0.0875	-0.4056
$k = 2$	0.2724	0.2784	-0.3622
$k = 3$	0.2638	0.1901	-0.3234
$k = 4$	0.2729	0.2069	-0.3424
$k = 5$	0.2700	0.2055	-0.3411

In addition to model sensitivity, we examined the rate of convergence of the recursive estimation procedure. Table 3 presents the results of the first five iterations.

5. Concluding remarks

In panel surveys where there is substantial overlap between monthly samples, state-space models enhanced with autoregressive structures can provide significant improvements over traditional design-based estimates of labor force characteristics, particularly in small areas. To be effective, the AR model should be carefully aligned with the rotation structure of the survey.

A central challenge in this context has been the estimation of the AR parameters, especially when they are embedded within a larger state-space framework. In this work, we have introduced a novel recursive algorithm that allows for the estimation of the AR coefficients from within the model itself—*without relying on external estimation procedures*. This integrated approach ensures consistency with the structure of the state-space model and enables estimation based solely on internal model dynamics.

We also proposed and demonstrated several diagnostic tools to assess the convergence, reliability, and sensitivity of the resulting estimates. These tools reinforce the empirical validity of the approach and support its potential for broader application in official statistics and beyond.

References

[1] Anderson, B.D.O., and Moore, J.B. (1979). *Optimal Filtering*. Prentice–Hall.

[2] Battese, G.E., Harter, R.M., and Fuller, W.A. (1988). An error-components model for prediction of county crop area using survey and satellite data. *Journal of the American Statistical Association*, 83, 28–36.

[3] Bertsekas, D.P. (1999). *Nonlinear Programming* (2nd ed.). Athena Scientific.

[4] Cochran, W.G. (1963). *Sampling Techniques* (2nd ed.). Wiley.

[5] Cochran, W.G. (1977). *Sampling Techniques* (3rd ed.). Wiley.

[6] Cohen, N., Ben-Hur, D., and Burck, L. (2017). Variance estimation in multi-phase calibration. *Survey Methodology*, 43(1), 125–140.

[7] Fay, R.E., and Herriot, R.A. (1979). Estimates of income for small places. *Journal of the American Statistical Association*, 74, 269–277.

[8] Fuller, W.A., and Rao, J.N.K. (2001). A regression composite estimator with application to Canadian Labor Force Survey. *Survey Methodology*, 27, 45–51.

[9] Harvey, A. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

[10] Horvitz, D., and Thompson, D. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47, 663–685.

[11] Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1), 35–45.
[12] Morris, N.D., and Pfeffermann, D. (1984). A Kalman filter approach to the forecasting of monthly time series affected by festivals. *Journal of Time Series Analysis*, 5, 255–268.
[13] Patterson, H.D. (1950). Sampling on successive occasions with partial replacement of units. *Journal of the Royal Statistical Society B*, 12, 241–255.
[14] Pfeffermann, D. (1991). Estimation and seasonal adjustment of population means using data from repeated surveys. *Journal of Business & Economic Statistics*, 9(2), 163–174.
[15] Pfeffermann, D., and Barnard, C.H. (1991). Some new estimators for small-area means with application to the assessment of farmland values. *Journal of Business & Economic Statistics*, 9(1), 73–84.
[16] Pfeffermann, D., and Burck, L. (1990). Robust small area estimation combining time series and cross-sectional data. *Survey Methodology*, 16(2), 217–237.
[17] Pfeffermann, D., Feder, M., and Signorelli, D. (1999). Estimation of autocorrelations of survey errors with application to trend estimation in small areas. *Journal of Business and Economic Statistics*, 16, 339–348.
[18] Pfeffermann, D., and Tiller, R.L. (2006). Small area estimation with state-space models subject to benchmark constraints. *Journal of the American Statistical Association*, 101, 1387–1397.
[19] Rao, J.N.K., and Graham, J.E. (1964). Rotation designs for sampling on repeated occasions. *Journal of the American Statistical Association*, 59, 492–509.
[20] Rao, J.N.K., and Yu, M. (1994). Small area estimation by combining time series and cross-sectional data. *Canadian Journal of Statistics*, 22, 511–528.
[21] Yates, F., and Grundy, P.M. (1953). Selection without replacement from within strata with probability proportional to size. *Journal of the Royal Statistical Society B*, 15, 235–261.
[22] Zhang, Q., and Yi, G.Y. (2022). Sensitivity analysis of error-contaminated time series data under autoregressive models with the application of COVID-19 data. *Journal of Applied Statistics*. doi: 10.1080/02664763.2022.2034760.